

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12Q 1/68	A2	(11) International Publication Number: WO 00/01846 (43) International Publication Date: 13 January 2000 (13.01.00)
(21) International Application Number: PCT/EP99/04718 (22) International Filing Date: 2 July 1999 (02.07.99) (30) Priority Data: 9814536.0 3 July 1998 (03.07.98) GB 9827152.1 9 December 1998 (09.12.98) GB (71) Applicant (for all designated States except US): DEVGEN N.V. [BE/BE]; Technologiepark 9, B-9052 Wondelgem (BE). (72) Inventors; and (75) Inventors/Applicants (for US only): PLAETINCK, Geert [BE/BE]; Devgen N.V., Technologiepark 9, B-9052 Wondelgem (BE). PLATTEEUW, Christ [BE/BE]; Devgen N.V., Technologiepark 9, B-9052 Wondelgem (BE). MORTIER, Katherine [BE/BE]; Devgen N.V., Technologiepark 9, B-9052 Wondelgem (BE). BOGAERT, Thierry [BE/BE]; Devgen N.V., Technologiepark 9, B-9052 Wondelgem (BE). (74) Agent: BOULT WADE TENNANT; 27 Fumival Street, London EC4A 1PQ (GB).		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: CHARACTERISATION OF GENE FUNCTION USING DOUBLE STRANDED RNA INHIBITION		
(57) Abstract <p>There is provided a method of identifying DNA responsible for conferring a particular phenotype in a cell which method comprises a) constructing a cDNA or genomic library of the DNA of said cell in a suitable vector in an orientation relative to a promoter(s) capable of initiating transcription of said cDNA or DNA to double stranded (ds) RNA upon binding of an appropriate transcription factor to said promoter(s), b) introducing said library into one or more of said cells comprising said transcription factor, and c) identifying and isolating a particular phenotype of said cell comprising said library and identifying the DNA or cDNA fragment from said library responsible for conferring said phenotype. Using this technique it is also possible to assign function to a known DNA sequence by a) identifying a homologue(s) of said DNA sequence in a cell, b) isolating the relevant DNA homologue(s) or a fragment thereof from said cell, c) cloning said homologue or fragment thereof into an appropriate vector in an orientation relative to a suitable promoter(s) capable of initiating transcription of dsRNA from said DNA homologue or fragment upon binding of an appropriate transcription factor to said promoter(s) and d) introducing said vector into said cell from step a) comprising said transcription factor.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

- 1 -

CHARACTERISATION OF GENE FUNCTION
USING DOUBLE STRANDED RNA INHIBITION

The present invention is concerned with
5 characterization or identification of gene function
using double stranded RNA inhibition (dsRNAi) and
methods of identifying DNA responsible for inducing a
specific phenotype in a cell and a method of assigning
function to known gene sequences.

10 It has recently been described in Nature Vol 391,
pp.806-811, February 98, that introducing double
stranded RNA into a cell results in potent and
specific interference with expression of endogenous
genes in the cell and which interference is
15 substantially more effective than providing either RNA
strand individually as proposed in antisense
technology. This specific reduction of the activity
of the gene was also found to occur in the nematode
worm *Caenorhabditis elegans* (*C. elegans*) when the RNA
20 was introduced into the genome or body cavity of the
worm.

The present inventors have utilized this
technique and applied it further to devise novel and
inventive methods of assigning functions to genes or
25 DNA fragments, which have been sequenced in various
projects, such as, for example, the human genome
project and which have yet to be accorded a particular
function and for use in identifying DNA responsible
for conferring a particular phenotype.

30 Therefore, according to a first aspect of the
present invention there is provided a method of
identifying DNA responsible for conferring a phenotype
in a cell which method comprises a) constructing a
cDNA or genomic library of the DNA of said cell in an
35 orientation relative to a promoter(s) capable of

promoting transcription of said cDNA or DNA to double stranded (ds) RNA upon binding of an appropriate transcription factor to said promoter(s), b) introducing said library into one or more of said cells comprising said transcription factor, and c) identifying and isolating a desired phenotype of said cell comprising said library and identifying the DNA or cDNA fragment from said library responsible for conferring said phenotype.

10 In a preferred embodiment of the invention the library may be organised into hierarchical pools as described in more detail in the examples provided, prior to step b) such as to include, for example, gene families.

15 According to a further aspect of the invention there is also provided a method of assigning function to a known DNA sequence which method comprises a) identifying a homologue(s) of said DNA in a cell, b) isolating the relevant DNA homologue(s) or a fragment thereof from said cell, c) cloning said homologue or
20 fragment into an appropriate vector in an orientation relative to a promoter(s) capable of promoting transcription of dsRNA upon binding of an appropriate transcription factor to said promoters, d) introducing said vector into said cell from step a) comprising
25 said transcription factor, and e) identifying the phenotype of said cell compared to wild type.

In each aspect of the invention, the nucleotide or DNA sequence may either be provided in a sense and
30 an antisense orientation relative to a single promoter which has the properties defined above, or alternatively it may be provided between two identical promoters. In both embodiments dsRNA is provided from the transcription initiated from the promoter
35 following binding of its appropriate transcription

factor.

The cell according to the invention may be derived from or contained in an organism. Where the cell is contained within an organism, the organism may be adapted to express the appropriate transcription factor. The organism may be any of a plant, animal, fungus or yeast but preferably may be the nematode worm *C. elegans*, which may be any of a wild type, a nuc-1 or pha-ts mutant of *C. elegans* or a combination of said mutations. In an alternative embodiment the DNA or cDNA library or the DNA homologue or fragment thereof may, advantageously, be transfected or transformed into a microorganism, such as a bacterial or yeast cell, which may be fed to the organism, which is preferably the nematode worm *C. elegans*. In this embodiment of the invention the microorganism may be adapted to express the appropriate transcription factor. Preferably, the microorganism is *E. coli*.

In each aspect of the invention, the DNA library, DNA homologue or DNA fragment may be constructed in a suitable DNA vector which comprises a sequence of nucleotides which encode said transcription factor. Alternatively, said transcription factor is encoded by a further vector. In an even further alternative, the cell or organism may express or be adapted to express said transcription factor. Preferably, any of the vectors used in the method according to the invention comprises a selectable marker which may be, for example, a nucleotide sequence encoding sup-35 or a fragment thereof. The nucleotide sequence may be orientated relative to a promoter such that binding of a transcription factor to the promoter initiates transcription of the DNA into double stranded RNA. Figure 10 illustrates the vectors and the orientation of the DNA sequence which enable double stranded RNA

production in *C. elegans*. Thus in one embodiment the DNA is located between two promoters on a vector capable of expressing dsRNA upon binding of an appropriate transcription factor to said promoters.

5 Alternatively, the vector comprises two copies of the DNA sequence organised in a sense and antisense orientation relative to the promoter and which marker is selectable when contained in a pha-1 mutant *C. elegans*. Preferably, the promoters are any of T7, T3

10 or SP6 promoters and the transcription factor comprises the appropriate polymerase.

Preferably, the selectable marker comprises a nucleotide sequence capable of inhibiting or preventing expression of a gene in said cell and which

15 gene is responsible for conferring a known phenotype. This nucleotide sequence may be part of or identical to said gene conferring said phenotype, and which nucleotide sequence is itself oriented relative to a suitable promoter(s) capable of initiating

20 transcription of double stranded RNA upon binding of an appropriate transcription factor to said promoter(s). Alternatively, the nucleotide sequence may be a part of or identical to said gene sequence conferring said phenotype, and which nucleotide

25 sequence is such as to permit integration of said suitable or further vector by homologous recombination in the genome of said cell and following said integration said nucleotide sequence is capable of inhibiting expression of said gene sequence conferring

30 said phenotype. In this embodiment said nucleotide sequence comprises stop codons sufficient to prevent translation of said nucleotide sequence following its integration into said genome.

Compounds can, advantageously, in said method be

35 added to said cell or organism for the purposes of

screening for desired phenotypes, such as for example, resistance or sensitivity to the compound when compared to wild type. The promoters are preferably inducible. The transcription factor may in some
5 embodiments be phage derived, such as for example, a T7 polymerase driven by a phage promoter. However, when *C. elegans* is utilised a worm specific or tissue specific promoter can be used, such as for example, let858, SERCA, UL6, myo-2 or myo-3. Preferably, the *E.*
10 coil strain is an RNAaseIII and even more preferably an Rnase negative strain.

A further aspect of the present invention provides a method of generating a transgenic non-human organism comprising an exogenous transcription factor
15 and a transgene comprising a promoter operably linked to DNA fragment which is expressed upon binding of said transcription factor thereto, the method comprising a) providing a first transgenic organism comprising a first construct incorporating DNA
20 encoding an exogenous transcription factor and a second transgenic organism comprising a second construct including at least one promoter operably linked to a desired DNA sequence which is expressed upon binding of the transcription factor of said first
25 transgenic organism thereto b) crossing said first and second transgenic organisms and selecting offspring expressing said desired DNA sequence. In one embodiment said first and second transgenic organisms are generated by transforming said first and second
30 constructs into respective microorganisms for subsequent feeding to the respective organism. Preferably, said second construct comprises said desired DNA sequence in an orientation relative to said promoter so as to be capable of initiating
35 transcription of said DNA to dsRNA upon binding of

- 6 -

said transcription factor thereto. In this embodiment said second construct comprises two promoters flanking said desired DNA sequence which promoters can initiate transcription of said DNA sequence to dsRNA upon
5 binding of said transcription factor to said promoters. Alternatively, said DNA sequence is provided in a sense and an antisense orientation relative to said promoter so as to produce dsRNA upon binding of the transcription factor to the promoters..
10 In each of these embodiments the first and/or second constructs may preferably be provided with a reporter gene operably linked to a promoter which is capable of initiating transcription of said reporter upon binding of said transcription factor thereto. Preferably, the
15 reporter gene encodes any of Luciferase, Green Fluorescent protein, β galactosidase or β -lactamase.

The present invention also includes a method of validating clones identified in yeast two hybrid vector experiments which experiments are well known to
20 those skilled in the art and which experiments were first proposed by Chien et al. (1991) to detect protein - protein interactions. The method according to the invention comprises providing a construct including the DNA encoding a protein identified in a
25 two hybrid vector experiment, which construct is such that said DNA is provided in an orientation relative to one or more promoters capable of promoting transcription of said DNA to double stranded RNA upon binding of an appropriate transcription factor to said
30 promoters, transforming a cell, such as a bacterial cell or alternatively transforming an organism comprising said transcription factor with said constructs and identifying a phenotypic change in said cell or organism, which may be *C. elegans* or the like,
35 compared to wild type. Preferably, the transcription

- 7 -

factor is inducible in the cell or organism. Once again the DNA sequence may be located between two promoters or in both a sense and antisense orientation relative to a single promoter, as described above.

5 Preferably, the promoter is a phage polymerase promoter and said transcription factor is a RNA polymerase, and preferably T7 polymerases. Also encompassed with the scope of the present invention are vectors used to transform said cells or organisms
10 and the cells or organisms themselves.

In a further aspect of the present invention there is provided a method of alleviating pest infestation of plants, which method comprises a) identifying a DNA sequence from said pest which is
15 critical either for its survival, growth, proliferation or reproduction, b) cloning said sequence from step a) or a fragment thereof in a suitable vector relative to one or more promoters capable of transcribing said sequence to RNA or dsRNA
20 upon binding of an appropriate transcription factor to said promoters, and c) introducing said vector into the plant.

Thus, advantageously, the method according to the invention provides a particularly selective mechanism
25 for alleviating pest infestation, and in some cases parasitic infestation of plants, such that when the pest feeds on the plant it will digest the expressed dsRNA in the plant thus inhibiting the expression of the DNA within the pest which is critical for its
30 growth, survival, proliferation or reproduction. In a preferred embodiment, the pest may be any of *Tylenchulus* ssp., *Radopholus* ssp., *Rhadinaphelenchus* ssp., *Heterodera* ssp., *Rotylenchulus* ssp., *Pratylenchus* ssp., *Belonolaimus* ssp., *Canjanus* ssp.,
35 *Meloidogyne* ssp., *Globodera* ssp., *Nacobbus* ssp.,

Ditylenchus ssp., Aphelenchoides ssp., Hirschmenniella ssp., Anguina ssp., Hoplolaimus ssp., Heliotylenchus ssp., Criconemellasp., Xiphinemassp., Longidorus ssp., Trichodorus ssp., Paratrichodorus ssp.,
5 Aphelenchs ssp. The DNA sequence or fragment thereof according to this aspect of the invention may be cloned between two tissue specific promoters, such as two root specific promoters.

A further aspect of the invention concerns the
10 vector used in each of the methods of the invention for constructing said library, which vector comprises two identical promoters oriented such that they are capable of initiating transcription of DNA sequence located between said promoters to dsRNA upon binding
15 of an appropriate transcription factor to said promoters. The DNA sequence may, for example, include a multiple cloning site. Preferably, the expression vector comprises a nucleotide sequence encoding a selectable marker. In one embodiment the nucleotide
20 sequence encoding said selectable marker is located between two identical promoters oriented such that they are capable of initiating transcription of DNA located between said promoters to double stranded RNA upon binding of an appropriate transcription factor to
25 said promoters. Preferably, the selectable marker comprises a nucleotide sequence encoding sup-35, for introduction into *C. elegans* having a pha-1 mutation.

Preferably, the transcription factor comprises either a phage polymerase which binds to its
30 corresponding promoter or a *C. elegans* specific promoter and even more preferably T7 polymerase. Preferably, the vector includes a multiple cloning site between said identical promoters.

In a further aspect of the invention there is
35 provided an expression vector for expressing an

appropriate transcription factor for use in a method according to the invention which vector comprises a sequence of nucleotides encoding said transcription factor operably linked to suitable expression control sequences. Preferably, the expression control sequences include promoters which are inducible, constitutive, general or tissue specific promoters, or combinations thereof. Preferably, the transcription factor comprises a phage polymerase, and preferably T7, T3 or SP6, RNA polymerase.

A further aspect of the invention provides a selection system for identifying transformation of a cell or organism with a vector according to the invention which system comprises a vector according to the invention wherein said selectable marker comprises a nucleotide sequence capable of inhibiting or preventing expression of a gene in said cell or organism which gene is responsible for conferring a known phenotype. Preferably said nucleotide sequence corresponds to a part of or is identical to said gene conferring said known phenotype, and which nucleotide sequence is itself located between two identical promoters capable of initiating transcription of double stranded RNA upon binding of an appropriate transcription factor thereto. Alternatively, the nucleotide sequence comprises a nucleotide sequence which is a part of or identical to said gene sequence which confers a known phenotype on said cell or organism, and which is such that following integration of said vector by homologous recombination in the chromosome of said cell or organism said sequence inhibits expression of said gene sequence conferring said known phenotype. Preferably, according to this embodiment the nucleotide sequence comprises stop codons sufficient

- 10 -

to prevent translation of the nucleotide sequence following integration into said chromosome.

Preferably, the known gene sequence comprises a sup-35 gene or a fragment thereof which is selectable by
5 identifying offspring growing at a temperature above 25°C following introduction in a pha-1 et123ts mutant *C. elegans* worm.

In a further aspect of the invention provides said known gene sequence comprises a sup-35 gene or a
10 fragment thereof which is selectable by identifying offspring growing at a temperature above 25°C following introduction of said vector in a pha-1 et123ts mutant *C. elegans* worm. An even further aspect comprises a method of assigning function to a
15 DNA sequence of a multicellular organism which method comprises a) providing i) a construct comprising said DNA fragment cloned between two promoters capable of promoting transcription in said multicellular organism, in a multicellular organism capable of
20 initiating transcription from said promoter; b) identifying the phenotype of said multicellular organism compared to wild type.

The present invention may be more clearly understood by the following examples which are purely
25 exemplary with reference to the accompanying figures, wherein:

Figure 1 is a nucleotide sequence of plasmid PGN1 in
30 accordance with the present invention.

Figure 2 is a nucleotide sequence of plasmid PGN100 in accordance with the present invention.

Figure 3 is a schematic representation of the vectors used and the transformation regime used in
35 the methods according to the present

- 11 -

invention.

Figure 4 is an illustration of an expression vector used in accordance with the invention.

5

Figure 5 is a schematic illustration of the T7 RNA polymerase expression vectors used for transforming *C. elegans*.

10 Figure 6 is an illustration of plasmid PGN1.

Figure 7 is a diagrammatic representation of an enhanced vector for dsRNA inhibition encoding sup-35 dsRNA.

15

Figure 8 is an illustration of a vector for integration into the genome of *C. elegans*.

20 Figure 9 is an illustration of the position of a DNA sequence(s) relative to a suitable promoter to initiate expression of dsRNA from the DNA sequence(s).

Figure 10 is a representation of plasmid pGN108.

25

Figure 11 is a representation of plasmid pGN105.

Figure 12 is a representation of plasmid pGN400.

30 Figure 13 is a representation of plasmid pGN401.

Figure 14 is a representation of plasmid pGN110.

35 Figure 15 is a representation of plasmid pAS2 with forward and reverse T7/T3/SP6 promoters.

- 12 -

Figure 16 is a representation of plasmid pGAD424 with forward and reverse T7/T3/SP6 promoters.

5 Figure 17 is a representation of plasmid pAS2-cyh2-HA+, both T7-final.

Figure 18 is a representation of plasmid pGAD424-without-FULL-ICE-BOTH-T7.

10 Figure 19 (a) is a representation of plasmid pGN205 and(b) is a representation of plasmid pGN207.

15 **Example A: Construction of an ordered and hierarchical pooled cDNA library and applications thereof.**

A random ordered and pooled library:

20 The vector is an *E. coli* vector harboring two T7 promoters, with a multiple cloning site (MCS) in between. The two promoters are orientated towards each other, and towards the MCS. In the presence of T7 RNA polymerase, expressed in *E. coli*, *C. elegans* or any other organism, RNA will be produced, starting
25 from the two T7 promoters. As these are oriented in the opposite sense, both strands of RNA will be produced from the DNA inserted (cloned) into the MCS in between the two promoters which results in the generation of double stranded RNA (dsRNA) upon binding
30 of the T7 RNA polymerase thereto.

A *C. elegans* cDNA library is constructed in the MCS using standard molecular biological techniques. The library is transformed into *E. coli*, and the resulting *E. coli* are grown in culture and stored in
35 96 multi-well plates. At this stage, plasmid DNA can

- 13 -

be isolated and stored in 96-multi-well plates corresponding to those of the *E. coli* colonies. Approximately 100,000 colonies are scored. In this way, the library will harbor approximately 5 times the total expressed cDNA variation of *C. elegans*, which gives the opportunity for low expressed sequences to be present in the library. This will result in approximately 1041 96-well plates. The plates are hierarchical pooled as necessary. For the present pooling of the clones is arranged in a range of 10 to 100. If the hierarchical pooling is per 8 or 12 (numbers are more convenient as 96-well plates have a 8 to 12 grid), this will result in approximately 87 multi-well plates and approximately 8352 wells. If hierarchical pooling is per 96 wells, which is a full plate, this results in approximately 11 plates and approximately 1041 wells. At any stage of the hierarchical pooling, plasmid DNA can be isolated, which would be less elaborate as less plates are used, but will result in a loss of complexity although this should not be the case in the pooling per 12. The pooling of the DNA can also be carried out with the original DNA.

The experiments below describe how the hierarchical pooling should be performed, both for the DNA and for the *E. coli* library.

An ordered library for RNAi technology, harboring every gene of the *C. elegans* genome, with applications thereof

As the genome-sequencing project is coming to an end, this information can be used in the application of T7 RNA inhibition technology. Every gene of the *C. elegans* genome can be cloned using PCR technology. In

- 14 -

preference, exons will be cloned with a minimal length of 500 bp. If the exons are too small, smaller fragments will be isolated with PCR, or even parts of introns and neighboring exons will be isolated with PCR technology so that at least a sufficient part of the translated region of the gene is cloned. For this, at least 17000 PCR reactions need to be performed. This collection of PCR products will be cloned in a T7 vector as described (two T7 promoters oriented towards each other with a multiple cloning site in between). Every PCR product is cloned independently, or can be used to generate a random library, analogous to the described cDNA library. If every PCR product is cloned individually, the resulting bacteria and plasmid DNA can be pooled in various ways. Firstly, this collection of individually cloned PCR products in the T7 RNAi vector can be pooled randomly, as described in the random library. This pooling can also be done in a more rational way. For instance, the genes of the *C. elegans* genome can be analyzed using bioinformatic tools (in silico biology). Various genes of the genome will belong to a gene family, or will have homologues in the genome. These members of the gene family will be pooled, or the members, being homologues will be pooled. In this way the total number of about 17000 clones is reduced to a more useable quantity. This library can be used to screen for phenotypes in the methods according to the invention. The resulting phenotype gives a functional description to the gene or gene family or gene homologues of the *C. elegans* genome. As the library consists of a part of every gene in the genome, this method enables description of the full genome in functional-phenotypic terms. For this the double

- 15 -

stranded RNA (dsRNA) needs to be introduced in the worm. This introduction of clones alone, or pooled clones, being random pooling or rational pooling can be achieved in several ways as described.

5

Example of a vector for the expression of double stranded RNAi

Any vector containing a T7 promoter may be used, and which contains a multiple cloning site (there are many commercially available). Primers containing the T7 promoter and a primer with the reverse complementary strand, both with the appropriate ends are designed. These primers can be hybridized, and if well designed, cloned in the vector of choice. The minimal sequence for a T7 promoter is TAATACGACTCACTATAGGGCGA. Although any vector can be used for the construction of a T7 expression vector there follows an example of how to achieve this with the vector pGEM-3zf(-).

20

- Vector pGEM-3zf(+) (PROMEGA) was digested with HindIII and SalI

25

- Primers oGN1 and oGN2 were mixed together at a final concentration of 1 μ g/30 μ l boiled and cooled slowly to room temperature.

30

- The primer was ligated into the vector using standard ligation procedures. The resulting vector is pGN1 (shown in Figure 1) and contains two T7 promoters oriented towards each other, and harbors a multiple cloning site in between.

35

Sequences of oGN1 and oGN2 are:

- oGN1: AGC TGT AAT ACG ACT CAC TAT AGG GCG AGA AGC TT
- oGN2: TCG AAA GCT TCT CGC ATA ATA GTG AGT CGT ATT AC

35

- 16 -

Example of the construction of a library

RNA may be isolated from every organism that is sensitive to RNAi. In general the isolated RNA is then copied into double stranded cDNA, and subsequently prepared in suitable vectors for cloning. Several procedures exist and molecular biology kits can be purchased from various firms including promega, clontech, boehringer Mannheim, BRL, etc which enable:

10

- isolation of RNA,
- eventually polyA RNA can be isolated (several techniques and kits available)
- first strand synthesis with AMV reverse transcriptase, random hexameric primers and/or oligo (dT) primer
- second strand synthesis with Rnase H, DNA PolymeraseI,
- flush ends with T4 DNA Polymerase
- addition of an adaptor with T4 DNA ligase.
- eventually treatment with T4 polynucleotide Kinase
- cloning of the cDNA into the vector.

15

20

The resulting ligation mixture can be considered as the cDNA library. The ligation contains all cDNA of the procedure ligated into the vector of interest. To order the library, the ligation needs to be transformed into *E. coli* strains.

25

30

Application of this *E. coli* or DNA library

T7 RNA producing strain:

- a standard strain is BL21 (DE3): F-ompT[lon]hsds(r-m-; and *E. coli* B strain) λ (DE3). Eventually variants of BL21 (DE3) can be used, although BL21

35

(DE3)pLysS is used.

- any other *E. coli* strain which produces the T7 RNA polymerase, which may be available needs to be constructed. This can be generated easily using a phage, which is commercially available, in this case, the λ CE6 vector (provided by Promega) is used. Almost every *E. coli* strain can be transfected with this phage and will produce T7 RNA polymerase.

- a RNaseIII mutant *E. coli*:

Various strains are in principle available, we chose in a first experiment to use strain AB301-105: rna-19, suc-11, bio-3, gdhA2, his95, rnc-105, relA1, spoT1, metB1. (Kinder et al. 1973 Mol. Gen. Genet 126:53), but other strains may suit better. This strain is infected with λ CE6 and so a T7 producing variant will be constructed.

Wild type *C. elegans* worms can be grown on the bacteria pools. The bacteria is expressing the T7 RNA polymerase. This results in large quantities of dsRNA in the gut of the *C. elegans*, which will diffuse in the organism and results in the inhibition of expression. This library can now be used for the screening of several phenotypes. This technique has the advantage that it is a much faster to detect relevant genes in certain pathways, than the known *C. elegans* technology. Moreover, if an interesting phenotype is found, the responsible gene can be cloned easily.

Using the hierarchical pooling one can easily find in a second screen the relevant clone of the pool. The inserted DNA of this clone can then be sequenced. This experiment results in genetic and biochemical DATA in one step.

Wild type *C. elegans* strains can be combined with compounds to screen for phenotype, drug resistance and

- 18 -

or drug sensibility. The *C. elegans* strain can be a mutant strain, screening for an enhanced phenotype, reduced phenotype, or a new phenotype. The *C. elegans* strain can be a mutant strain, and the library screen can be combined with compounds. So one can screen for drug resistance, drug sensibility, enhanced phenotype, reduced phenotype, or a new phenotype. The *E. coli* strain may be any T7 RNA polymerase expressing strain, like BL21 (DE3), for example, but the formation of double strand RNA may be enhanced by using a special *E. coli* strain that is RNaseIII negative. RNaseIII recognizes specific loops in dsRNA. Eventually, an *E. coli* strain can be used that is deleted in RNases other than RNaseIII or an *E. coli* can be used that is deleted in one or more RNases. The expression of the T7 RNA polymerase in most known *E. coli* strains and constructs which are available to generate T7 RNA polymerase producing *E. coli* strains, generally comprise an inducible promoter. In this way the production of the T7 RNA polymerase is regulated, and thus the production of the dsRNA. Advantageously, this feature can be used to "pulse" feed the *C. elegans* worms at specific stages of growth. The worms are grown on the non-induced *E. coli* strains. When the worm has reached the stage of interest, the T7 RNA production in the bacteria is induced. This allows the studying of the function of any gene at any point in the life cycle of the animal.

Screening the library for homologues of putative interesting human genes, and assign function to these genes

Hundreds of genes have been isolated in various projects, being genomic projects, differential

- 19 -

expressed arrays, hybridization studies, etc. The described cDNA library can provide a way to validate and or assign function to these genes in a fast and efficient manner. First of all the worm homologue or homologues or the genes need to be identified by bioinformatic tools (in silico biology). PCR primers are developed and the cDNA fragment is isolated using PCR technology. PCR can be performed on the hierarchical pools. The positive pool or individual wells harboring the bacteria that has the appropriate cDNA is fed to *C. elegans* and the phenotype is scored.

PCR can be performed on cDNA isolated from *C. elegans*. The resulting DNA can be cloned in the T7 vector and transformed in the dsRNA producing *E. coli* on which the *C. elegans* worms are then fed. Depending on which way is faster and more reliable a choice needs to be made.

If the gene belongs to a gene family, the worm may need to be fed on a mixture of bacteria. Each of them harboring a part of the member of the gene family. *E. coli* strains, growth conditions, combinations with compounds can be performed as described above.

If the library rational is used, in which all the genes of *C. elegans* are cloned in a organized and structured way, the *C. elegans* homologue and eventually the other homologues, orthologues, and members of the gene family can be traced back easily in the library using a silico biology. No PCR is involved in this step, and the bacteria and or DNA can be isolated on which the worm will be grown.

Examples

The idea of the series of experiments was to test

- 20 -

both the RNAi vector and the various *E. coli* strains that were constructed.

1) Construction of a test plasmid

5 Any cDNA that gives a clear phenotype in the worm when knocked-out, or used in a RNAi experiment can be used. It is known that *unc-22* is a good candidate, but many other genes are possible. We opted for a sensitive system that can be used at a later stage.

10 The system was tested with *sup-35* in a *pha-1* background. Exon 5 of the *sup-35* was isolated by PCR and cloned in the T7 promoter vector pGN1. The resulting vector was designated pGN2. *pha-1* (e2123) mutant worms cannot produce offspring at temperatures

15 higher than 25°C. This is due to a developmental problem in embryogenesis. When *sup-35* is knocked-out, or inhibited in this strain, offspring may grow at this temperature. Combination of *pha-1* mutant worms and *sup-35* RNAi is a good system to validate the

20 various options.

2) Testing the RNAi using an *E. coli* strain that produces dsRNA.

- pGN2 was introduced in *E. coli* strain BL21(DE3) and

25 T7 RNA polymerase was induced with IPTG. *C. elegans* worms (*pha-1* (e2123)) were inoculated on this bacteria, and grown at the restricted temperature of 25°C. As this mutant is an embryonic mutant at this temperature, no offspring will be observed. If the

30 *sup-35* gene is efficiently inhibited by the dsRNA present in the *E. coli*, offspring will be observed.

- pGN2 was introduced in *E. coli* strain AB301-105(DE3) and T7 RNA polymerase was induced with IPTG. *C. elegans* worms (*pha-1* (e2123)) were inoculated on this

35 bacteria, and grown at the restricted temperature of

- 21 -

25°C. As this mutant is an embryonic mutant at this temperature, no offspring will be observed. If the sup-35 gene is efficiently inhibited by the dsRNA present in the *E. coli*, offspring will be observed.

5

3) Improving the worm strain for better uptake of dsRNA.

Before plating the pha-1 *C. elegans* on the *E. coli* strain that produce the double stranded sup-35 RNA. The worm was mutagenised with EMS (Methane sulfonic Acid Ethyl). The offspring of this mutagenised worm is then plated on the bacteria. The worm that feed on this bacteria give larger offspring which has a mutation that results in an improvement of dsRNA uptake, and can be used for further experiments.

15

Stable integration of the dsRNA producing vector into the genome of the T7 RNA polymerase producing worm

20

An *E. coli* vector can be constructed harboring the following features; Two T7 promoters directed towards each other, with a restriction site or a multiple cloning site in between. Furthermore, the vector may contain the *C. elegans* sup35 genomic DNA, engineered in such a way that it contains several stopcodons at various intervals, so that no full length protein can be expressed from the sup35 genomic DNA fragment as illustrated in Figure 8. Any cDNA or cDNA fragment can be cloned in the multiple cloning site between the two T7 promoters. When this vector is introduced in a *C. elegans* strain which expresses T7 RNA polymerase, the cDNA or DNA fragment cloned between the two T7 promoters will be transcribed, generating dsRNA from the cloned fragment.

35

- 22 -

The vector is designed to be used in pha-1 (e2123) mutant worms expressing T7 RNA polymerase. The expression of the T7 RNA polymerase may be constitutive or regulated, general or tissue specific. These pha-1 (e2123) worms cannot produce offspring at temperatures higher than 25°C, which is due to a development problem in embryogenesis. When sup-35 is inhibited or knocked-out in this strain, offspring may grow at this temperature.

When the vector is introduced in the worm, the vector may integrate by homologous recombination (Campbell-like integration). It has been shown that homologous recombination occurs in *C. elegans*, although at low frequencies (Plasterk and Groenen, EMBO J. 11:287-290, 1992). Homologous recombination at the sup35 gene will result in a knock-out of the gene as the two resulting sup-35 genes will harbor the stopcodons. The resulting worm, and its offspring, if this recombination happens in the eggs, will have a copy of the vector integrated in the genome. This can be selected as only the worms for which the sup-35 has been knocked-out will have offspring at temperatures higher than 25°C. Furthermore, the resulting worm will stably produce double stranded RNA from the DNA fragment cloned between the two T7 promoters. This worm can now be considered as a stable transgenic worm strain with a reduction of function of the gene, from which a fragment has been cloned between the two T7 promoters.

The DNA may be delivered to the worm by several techniques, including injection, ballistic transformation, soaking in the DNA solution, feeding with bacteria. New and other methods that increase the transformation efficiencies can be considered.

The target *C. elegans* strain may in addition,

have other mutations than the pha-1 (e2123) mutation, and may express other genes than T7 RNA polymerase.

EXAMPLE B: a Yeast two-hybrid-RNAi vector

5
A yeast two hybrid vector can be constructed harboring the two T7 promoters. The vectors can be designed to replicate both in yeast and in *E. coil*. In general cDNA libraries for the yeast two hybrid
10 system are made in the Gal4 or LexA vectors. The library is constructed in vectors having the activation domain of one of these genes. A vector can be constructed that can still perform in the yeast two hybrid screen but which also contains two T7
15 promoters orientated towards each other, with a cloning site therein between. The order of the sequences in the plasmid will then be "plasmid backbone, (GAL4-T7), MCS, T7, backbone". A *C. elegans* cDNA library constructed in this vector can be used as
20 a standard yeast two hybrid library in an experiment to isolate interacting proteins with a given protein. Once a clone is isolated, the plasmid can be introduced in an *E. coil* strain expressing the T7 RNA polymerase, and hence will produce dsRNA of the
25 cloned fragment. The bacteria producing this dsRNA can be fed to the worm and phenotypes can be scored. As in the previous example, this validation procedure for a newly isolated yeast two hybrid clone is remarkably shorter than the standard procedure, which
30 requires PCR and/or cloning steps, RNA experiments and/or knock-out experiments. In most cases isolated clones are sequenced first, and on the basis of the sequence, a decision is made to continue with further experiments. In the present invention every isolated
35 clone can easily be introduced into the appropriate *E.*

- 24 -

coli and fed to the worm. Validation is then performed by phenotype analysis.

To apply this procedure a yeast two hybrid was performed using a known gene as bait and the newly constructed library as the target. Proteins coded by the clones in the target that interact with the bait protein, will result in positive yeast clones expressing the reporter molecule such as can be observed by LacZ staining with X-gal. The plasmid coding for the target protein is isolated directly from the yeast strain and introduced in *E. coli*. The *E. coli* is T7 RNA polymerase producing *E. coli*. In this case, double stranded RNA is produced from the DNA cloned in the multiple cloning site of the vector. When this dsRNA is fed to the worm using the methods described previously, the gene has inhibited in the worm, resulting in a particular phenotype.

- This yeast two hybrid vector can advantageously be used to construct an ordered and hierarchically pooled library as described in the previous example.
- A yeast strain can also be constructed that conditionally produces T7 RNA polymerase. After yeast two hybrid experiments, the expression of the T7 polymerase could be induced, resulting in the production of dsRNA in the yeast cell. Consequently the yeast could be fed to the worm. Evidence is available showing that the *C. elegans* worms can feed on yeast.

Construction of a T7 RNA polymerase producing strain, and applications thereof

A *C. elegans* strain can be constructed that expresses T7 RNA polymerase. The expression can be general and constitutive, but could also be regulated

- 25 -

under a tissue specific promoter, an inducible promoter, or a temporal promoter or a promoter that harbors one of these characteristics or combination of characteristics. DNA can be introduced in this *C. elegans* strain. This is done either by injection, by shooting with particles, by electroporation or as aforementioned by feeding. If the DNA is a plasmid as described in the previous examples, i.e. a plasmid harboring a cloned cDNA fragment or a PCR fragment between two flanking T7 promoters, then dsRNA of this cDNA or PCR fragment is formed in the cell or whole organism resulting in down regulation of the corresponding gene. The introduced DNA can have an efficient transient down regulation. The introduced DNA can form an extrachromosomal array, which array might result in a more catalytic knock-out or reduction of function phenotype. The plasmid might also integrate into the genome of the organism, resulting in the same catalytic knock out or reduction of function phenotype, but which is stably transmittable.

- Plasmid DNA harboring a cDNA or a part of a cDNA or an EST or an PCR fragment of *C. elegans* cloned between two T7 promoters as described in Examples A) and B) can be introduced in the T7 RNA polymerase worm, by standard techniques. Phenotypes can be analysed -DNA from an ordered and pooled library as in Example A) can be introduced in the T7 RNA polymerase worm, by standard techniques (injection, shooting). Phenotypes can be analysed. With the hierarchical pool, the original clone can be found easily.
- The same procedure can be performed with a mutant worm expressing the T7 RNA polymerase. Screening for enhanced, reduced or new phenotypes.
- The procedure can be used to enable screening of

- 26 -

compounds. Screening with either a wild-type strain or a mutant strain for enhanced or new phenotypes.

- The DNA could be introduced in the worm by new methods. One of which is the delivery of DNA by *E. coli*. In this case the hierarchical pooled library is fed to the animal. To prevent digestion of the *E. coli* DNA in the gut of the nematode, preferentially a DNase deficient *C. elegans* will be used, such as nuc-1 (el392). This procedure would be one of the most interesting as it would be independent of transformation efficiencies of other techniques, and generally faster and less labourious.

2) Putative enhancements of the method.

- A vector is designed, so that it harbors the sup-35 cDNA or a part of this cDNA, cloned in between two T7 promoters. The rest of the vector is as described in Examples A) and B). This vector can be introduced into a pha-1ts mutant *C. elegans*. A temperature selection system exists in this case and only those worms which have taken up the DNA and express the double stranded sup-35 RNA will survive at restricted temperatures. The hierarchical pooled library can be delivered by any method described above.
- The vector can be used to construct a library that is introduced in a T7 RNA polymerase expressing *E. coli*. In this case we have an analogous screening as in part A) with an additional screening for worms where the dsRNA of sup-35 is active.
- The DNA and or dsRNA of sup-35 could be delivered on a different plasmid. For the feeding, both DNA feeding (Example C) or dsRNA feeding Example A) and B), this means that the two plasmids could be present in one bacterium, or that the worm is fed on a mixture of bacteria, one of which harbors the sup-35 construct.

- 27 -

**Example of the construction of a T7 RNA producing
C. elegans**

To produce T7 RNA polymerase in the worm, several possibilities are possible. The T7 polymerase can be expressed under various promoters, being inducible promoters, constitutive promoters, general promoters and tissue (cell) specific promoters, or combinations of those. Examples of these promoters are the heatshock promoter hsp-16, the gut promoter ges 1, the promoter from cet858, but also the promoter of dpy 7 and the promoter element GATA1. In this example the T7 RNA polymerase is expressed under the control of the hsp-16 promoter that is available in the pPD49.78 vector. The T7 RNA polymerase is isolated as a PCR product using the primers of GN3 and GN4.

The resulting PCR product is digested with NheI and NcoI, as is the vector in which we want to clone, being the Fire vector pPD49.78. The resulting vector is pGN100 illustrated in Figure 2. oGN3: CAT GGC AGG ATG AAC ACG ATT AAC ATC GC oGN4: ATG GCC CCA TGG TTA CGG GAA CGC GAA GTC CG pGN100 is included.

The vector is introduced into the worm using standard techniques, such as micro injection, for example.

The following strains were then constructed:

- Wild-type (pGN100)
- nuc-1 (e1392) (pGN100)
- pha-1 (e2123) (pGN100)
- pha-1; nuc-1 (pGN100)

All of these strains are able to produce T7 RNA polymerase when temperature induced or alternatively

- 28 -

by metals such as application of heavy cadmium or mercury. The procedure for temperature induction is to shift the animal to a temperature of 30-33°C for at least one hour, then the animal can be shifted back to standard
5 temperatures (15-25°C).

The wild type strain producing T7 RNA polymerase can be used for the production of any RNA in the worm. More specifically, the plasmids from the described
10 libraries can be introduced in these worms, and phenotypes can be scored.

The nuc-1 mutant worm will be used to introduce DNA via bacteria on which the worm feed. As the nuc-1 worm does not digest the DNA, the plasmid DNA can
15 cross the gut wall. If taken up by the cells that produce the T7 RNA polymerase, dsRNA will be produced thus inhibiting the gene from which the RNA was transcribed.

The pha-1 mutant strain that produced T7 RNA polymerase can be used to enhance the procedures as
20 described above. DNA can be introduced by shooting, micro injection or feeding. More specifically this strain can be used for the vectors that produce dsRNA from sup-35 and from the gene of interest, the latter
25 can be a PCR product, a cDNA, or a library as described.

The pha-1; nuc-1 mutant producing T7 RNA polymerase can be used for the bacterial delivery of the DNA. DNA will preferentially be the plasmid that
30 produce dsRNA from both sup-35 and the gene of interest. The worm strain will preferentially produce the T7 RNA polymerase in the gut. Delivery will preferentially happen by feeding the worm on bacteria harboring the plasmid.

35

Application of the RNAi technology in plants

Nematodes are responsible a large part of the damage inflicted on plants and more particularly to plants used in the agricultural industry. The RNAi procedures according to the invention can be applied to plants to prevent these parasitic nematodes from feeding longer. In a first step, a DNA fragment is isolated from the parasitic plant nematode that is critical for the animals survival or growth, or to feed or to proliferate. Any gene from which the expression is essential is suitable for this purpose.

A part of this gene, an exon or cDNA is cloned. This DNA fragment can be cloned under the influence of a tissue specific promoter preferably a root specific promoter even more preferably between two root specific promoters. The DNA of the cloned gene under the control of the root specific promoter can be introduced in the plant of interest, using plant transgenic technology. For every parasitic nematode, a different piece of DNA may be required and likewise for every plant race, a different promoter will be needed.

The root will produce RNA or dsRNA from the introduced piece of DNA when root specific promoter is utilised. As the nematode feeds on the plant, the RNA and/or dsRNA will be consumed or ingested by the nematode. The RNA and/or dsRNA can enter the cells of the nematode and perform its inhibitory action on the target DNA. Depending on the nature of the cloned DNA piece of worm, the nematode will not be able to survive, to eat, proliferate, etc in any case preventing the animal of feeding longer on the plant, and thus protecting the plant.

- 30 -

Construction of a T7 RNA polymerase producing *C. elegans*

To produce a T7 RNA polymerase or other RNA
5 polymerases in animals, and more particularly in
nematodes and most particularly in *C. elegans*, several
possibilities can be envisaged. The T7 RNA polymerase
can be expressed under various promoters. These
promoters may be inducible promoters, constitutive
10 promoters, general promoters, tissue specific
promoters, or combinations of those.

Example 1:

Construction of a basic vector for expression of
15 T7 polymerase in *C. elegans*

The T7 polymerase coding sequence was PCR
amplified from λ CE6 (Novagen, Madison, USA) using the
primers oGN26 (ATGGAATTCTTACGCGAACGCGAAGTCCG) and
20 oGN46 (CTCACCGGTAATGAACACGATTAAACATCGC), using standard
procedures (PCR, A practical A practical approach,
1993, Ed. J. McPherson, et al, IRL Press). The
resulting DNA fragment encoding for the T7 RNA
polymerase was digested with AgeI and EcoRI and
25 inserted into the Fire vector pPD97.82 digested with
AgeI and EcoRI. The resulting construct encodes for an
open reading frame of T7 RNA polymerase in fusion with
the SV40 large T antigen nuclear localization signal
(NLS) with amino acid sequence MTAPKKRKVPV. This
30 nuclear localization signal sequence is required to
translocate the T7 RNA polymerase from the cytoplasm
to the nucleus, where it is able to bind to its
specific promoters, designated T7 promoters. Upstream
of the coding sequence for the T7 polymerase fusion
35 protein is a minimal promoter (*myo-2*) preceded by a

- 31 -

multiple cloning site (MCS) in which several C. elegans promoters can be inserted. This plasmid (pGN105 shown in Figure 11) is a basic T7 RNA polymerase plasmid which enables the expression of T7 polymerase in C. elegans. Derivatives of this plasmid wherein promoters are cloned into the multiple cloning site, allow for the inducible, constitutive, general and tissue specific expression of T7 RNA polymerase in C. elegans, as expression will be regulated by the promoter cloned in the multiple cloning site.

Although not restricted to these examples, for the following promoters it is known that they induce expression in the following tissues.

let-858 (ubiquitous expression), myo-2 (pharynx expression), myo-3 (body wall muscles), egl-15 (vulval muscles), unc-119 (pan-neuron).

Example 2:

Construction of a vector for expression of T7 RNA polymerase in C. elegans muscle tissue.

The T7 RNA polymerase coding sequence was PCR amplified from λ CE6 using the primers oGN43 (GCCACCGGTGCGAGCTCATGAACACGATTAACATCGC) and oGN44 (CACTAGTGGGCCCTTACGCGAACGCGAAGTCCG) digested with AgeI/SpeI and inserted in the pGK13 vector digested with AgeI/SpeI. (This vector contains the strong SERCA promoter which drives expression in the pharynx, the vulval muscle, the tail and the body wall muscle). A nuclear localization signal (NLS) of SV40 large T antigen was inserted in front of the T7 polymerase coding sequence by insertion of two overlapping oligo's oGN45 (CCGGATGACTGCTCCAAAGAAGAAGCGTAAGCT) and oGN46 (CTCACCGGTAATGAACACGATTAACATCGC) into the

- 32 -

SacI/AgeI restriction sites. The resulting construct was called pGN108 as shown in Figure 10. Introduction of this plasmid into *C. elegans* results in the expression of T7 RNA polymerase in the pharynx, vulva muscle, tail and body wall muscles.

To test expression and functionality of T7 RNA polymerase in *C. elegans* under the regulation of the SERCA promoter, pGN108, which encodes the T7RNA polymerase under the control of the SERCA promoter was injected into *C. elegans*. A test vector was co-injected. This test vector encodes for GFP under the control of a T7 promoter (pGN401 in Figure 13). The plasmid pGN401 was constructed by inserting two overlapping oligo's oGN41 (CCCGGGATTAATACGACTCACTATA) and oGN42 (CCGGTATAGTGAGTCGTATTAAATCCCGGGAGCT) in the SacI/AgeI opened Fire vector pPD97.82. generating a T7 promoter. Furthermore a selection marker was co-injected to select for transformants (rol6, pRF4). The latter selection vector pRF4 is well known to person skilled in the art. Transgenic F1 could easily be isolated as they display the rol 6 phenotype. These transgenic *C. elegans* all expressed GFP in the pharynx, the vulval muscle, the tail and the body wall muscle. This data show clearly that the T7 RNA polymerase is functionally expressed under the regulation of the SERCA promoter, and that the expressed T7 RNA polymerase binds to the T7 promoter present in pGN401 and initiates transcription of the GFP gene, which is then functionally expressed, resulting in fluorescence in the muscle tissues where SERCA is inducing the expression of the T7 RNA polymerase.

Example 3:Construction of a vector for ubiquitous
expression of T7 polymerase in C. elegans

5 The NLS-T7 RNA polymerase fusion gene was
isolated from pGN108 with XmaI/Bsp120I and cloned into
the Fire vector pPD103.05 digested with
XmaI/Bsp120I. This results in a vector wherein the T7
RNA polymerase is cloned under the regulation of the
10 let858 promoter. This specific promoter enables the
expression of T7 RNA polymerase in all tissues. The
resulting plasmid was named pGN110 (Figure 14).

Example 4:

15 Construction of a vector for T7 RNA polymerase
mediated expression of DNA fragments, genes, and
cDNA's under the control of a T7 promoter.

20 The Fire vector pPD97.82 was digested with SacI /
AgeI and a T7 promoter sequence was generated by
insertion of two overlapping oligo's oGN41
(CCCGGGATTAATACGACTCACTATA) and oGN42
(CCGGTATAGTGAGTCGTATTAATCCCGGAGCT) into the SacI/AgeI/
restriction endonuclease sites. This construct (pGN400
25 Figure 12) contains a GFP open reading frame cloned
between SacI and EcoRI restriction endonuclease sites
under the regulation of the T7 promoter. Any gene,
cDNA, or DNA fragment can be cloned in this vector by
deleting the GFP gene as a AgeI/ SacI fragment and
30 cloning the DNA fragment of interest into the vector.
Preferentially the DNA fragment of interest can be
obtained by PCR amplification, inserting the SacI/AfeI
sites in the primers. The resulting DNA fragment after
PCR amplification is the digested and the GFP gene in
35 pGN400 is replaced by the amplified DNA fragment.

- 34 -

Every vector that contains a T7 promoter could be used for the purpose of T7 RNA polymerase induced expression in *C. elegans*, such as the commercially available pGEM vectors and the pBluescript vectors.

5 This is clearly shown by the pGN401 vector which expresses GFP under the regulation of the T7 promoter in a transgenic *C. elegans* which expresses T7 RNA polymerase.

10 The use of pGN400 has the advantage that the vector includes a 3'UTR fragment from unc-54 which enhances the transcription or stability of the RNA.

Generation of permanent, tissue specific "pseudo knock-out" RNAi *C. elegans* lines

15

At present, gene knock outs in *C. elegans* are obtained after random, large scale mutagenesis and PCR base sib-selection. This method is bulky, very time consuming and tedious. It has been described that
20 introducing double stranded RNA into a cell results in potent and specific interference of expression of endogenous genes. In *C. elegans* gene expression can be down regulated by injection of RNA into the body cavity of the worm, soaking the worm in a solution
25 containing dsRNA or feeding *E. coli* that express dsRNA corresponding to the gene of interest. *C. elegans* cells have the ability to take in dsRNA from their extracellular environment. It has been reported that mRNA is the target of this ds RNA mediated genetic
30 interference (Montgomery and Fire 1998). It is also suggested that the targeted RNA is degraded in the nucleus before translation can occur. Although the RNAi mediated reduction of gene expression can be passed on to the next generations, heritability is
35 poor and the effect is rapidly lost during further

- 35 -

offspring. This is probably due to a continued decrease of the dsRNA pool. We propose here a method to construct *C. elegans* lines with a permanent, inheritable, RNAi phenotype. The method encompasses the generation of transgenic *C. elegans* lines by introducing plasmids containing cDNA fragments of the target gene in the sense and antisense orientation under control of a worm promoter or by transcription of an inverted repeat of the cDNA from a single construct. Alternatively, ds RNA can be transcribed from a vector harboring a cDNA flanked by two T7 promoters in a *C. elegans* strain that expresses T7 polymerase. The result is a transgenic worm with an heritable stable "pseudo knock-out" phenotype. The expression of the cDNA or the T7 polymerase can be general and constitutive but could also be regulated under a tissue specific promoter. In contrast to RNAi induced by external ds RNAi (injected, soaked or fed) this method would enable to obtain conditional, tissue specific inhibition of gene expression.

Inhibition of unc-22 expression by RNA interference results in a "twitching" phenotype.

25

Unc 22 cDNA (exon 22) was cloned in sense and antisense orientation in pPD103.05. (A. Fire nr L2865) containing the let 858 promoter that is capable of expressing RNA sequences in all tissues. The resulting plasmids were named pGN205 (Figure 19a) and pGN207 (Figure 19 b). These constructs were introduced into *C. elegans* together with a selectable marker (rol-6; GFP). Transgenic F1 individuals (expressing rol-6 or GFP) showed a "twitching" phenotype indicating that RNAi could be mediated by endogenous transcription of

- 36 -

RNA from transgenic DNA. The RNAi phenotype co-segregated with the selectable marker into further offspring. This resulted in the generation of *C. elegans* lines with permanent RNAi phenotype.

5

Generation of stable lines T7 RNA polymerase lines and generation of dual transgenic worms.

10 An expression system in *C. elegans* based on an exogenous RNA polymerase demands two plasmids. One is encoded for the RNA polymerase under the control of a specific promoter, while the other plasmid encodes for the DNA fragment to be expressed, under the regulation of the T7 promoter. In the case of semi stable RNAi
15 also designated pseudo stable knockouts, the DNA of interest is cloned between two T7 promoters so that dsRNA can be produced.

As the T7 RNA polymerase expression system is known to be a high expression system this will result
20 in problems to generate dually transgenic animals. If the gene to be expressed in the *C. elegans* nematode is toxic, this will result in lethal effects and hence in the construction of a *C. elegans* without highly regulated stable expression of the gene of interest.
25 If the gene of interest is essential for the survival of the organism, RNAi with a DNA fragment from this gene will also result in lethal effects, so that pseudo- stable knockouts are not possible.

To overcome this problem the present inventors
30 have designed a system consisting of two transgenic animals. The first animal is transgenic for the T7 RNA polymerase, This T7 RNA polymerase can be expressed in all cells or specific cells or tissues as has been shown in previous examples. The second transgenic
35 animal is transgenic for the DNA fragment of interest.

- 37 -

This can be a gene or cDNA linked to a T7 promoter, or if one wants to perform RNAi a DNA fragment of such gene cloned between two T7 promoters.

Both transgenic animals are viable and do not show any aberrant phenotypes. This is because the T7RNA polymerase expressed in the first transgenic organism is not toxic for the organism, even if expressed at relative high levels. In the second transgenic organism, the gene of interest is not expressed or the dsRNA is not produced as these transgenic animals do not contain the T7 RNA polymerase.

Expression of the gene or cDNA of interest or RNAi with a DNA fragment can now be obtained by mating the two transgenic animals. The offspring of these are dually transgenic and express the gene of interest or express dsRNA of the DNA fragment of interest. To generate sufficient males in such a mating, one of the transgenic animals males can be a *C. elegans* mutant with a phenotype favouring generation of males. An Example of such a mutant is him-5. Preferentially such a mutant will be used to make a *C. elegans* transgenic for T7 RNA polymerase, while the hermaphrodite harbors the DNA fragment under the regulation of the T7 promoter.

To select efficiently for the dual transgenic offspring a second transgene can be introduced in the second transgenic animal. This transgene contains a reporter gene under the regulation of the T7 promoter. The reporter gene can be GFP, luciferase, Beta galactosidase, or beta-lactamase. an example of such a ttransgene are the vectors pGN400 and pGN401.

To obtain inducible, tissue specific expression of a transgene in *C. elegans* we can make male stock (i.e. him-5) carrying the T7 polymerase construct

under the control of different *C. elegans* promoters that enable tissue specific expression such as). This males can be crossed with hermaphrodites carrying the gene of interest under the control of a T7 promoter.

5 Furthermore, the transgenes can be integrated into the genome of the animal. Methods to generate stable integration of a plasmid into the genome of the animal have been described (Methods in cell biology, Vol. 48, 1995, ed. by Epstein and Shakes, academic
10 press) and involve radiation of the animal. This can be done for both animals animals, but preferentially, the animals expressing the T7 RNA polymerase are subject to such treatment. This result in a collection of *C. elegans* nematodes that stably
15 express T7 RNA polymerase under the control of various promoters. examples of such promoters are the *myo-2* (pharynx expression), *myo-3* (body wall muscles), *egl-15* (vulval muscles), *unc-119* (pan-neuron), *SERCA* (muscles), *let858* (all cells) *ges-1* (gut).

20

Construction of RNAi T7 promoter yeast two hybrid vectors

pGAD424 with forward and reverse T7/T3 and or Sp6

25 In most two-hybrid experiments a cDNA library is cloned in plasmid pGAD424 (Figure 16) which has been engineered with additional restriction sites in the polylinker such as a *NcoI* site (Clontech). This library allows for screening of binding proteins in a
30 yeast two hybrid experiment. We constructed a new yeast two hybrid vector with the same possibilities to perform yeast two hybrid, but which contain two additional T7 promoters, so that the vector can be used for T7 RNA polymerase induced pseudo-stable
35 knock-outs. For this we inserted a forward T7 by using

- 39 -

a T7- linker (consisting of the following primers
aattccttaatacgaactcactatagggcc and
catggggccctatagtgagtcgtattaag) into the EcoRI-NcoI site
of pGAD424. The resulting vector was designated
5 pGAD424-without-FULL-ICE-both-T7. Care was taken to
eliminate stop codons and using maximal polylinker
compatible amino acids. We adopted the same strategy
for the reverse T7 (consisting of both primers
gatccgtcgacagatctccctatagtgagtcgtattactgca and
10 gtaatacgaactcactatagggagatctgtcgacg) with BamHI and
PstI. To avoid loss of SalI, we included this site in
the primer.

The SalI site is important as most libraries are
cloned in this site, adapters are available. This
15 makes the newly constructed vector compatible with
existing vectors.

pAS2 with with forward and reverse T7/T3 and or
Sp6

20 An analogous yeast two hybrid vector was
constructed based on pAS2 (Clontech). By partial EcoRV
digestion we were able to remove a significant part of
the cyh2 gene. The right construct can be isolated and
checked by a restriction digest with BglII. this
25 restriction site is present in the EcoRV fragment of
PAS2 to be eliminated. This eliminates the cyh2 gene
which is slightly toxic gene and involved in growth
retardation. This gene is non-essential for the
performing of RNAi and Yeast two hybrid experiments.
30 After the elimination of the EcoRV fragment, The
EcoRI restriction site which is located between the
DNA sequence encoding for GAL4DB and HA (epitope)
becomes unique for the plasmid, and can be used to
35 substitute HA with a T7 promoter containing linker.

- 40 -

This ensures persistence of all restriction sites, allowing both in frame cloning and compatibility with previous vectors and pGAD424. We used the following linker (primers: aattcttaatacgcactcactatagggca and
5 tatgccctatagtgagtcgtattaag) using EcorI and NdeI cloning sites. We adopted the same strategy for the reverse T7 (primers:
gatccgtcgacagatctccctatagtgagtcgtattactgca
catgggccctatagtgagtcgtattaag and
10 gtaatacgcactcactatagggagatctgtcgacg) with BamHI and PstI. To avoid loss of SalI we included it in the primer. The resulting vector was designated pAS2-cyh2-HA+both T7-final.

Having the T7 promoter (or alternatively the T3, or SP6 promoter) in pGAD424 allows to go quickly from
15 interacting protein to RNAi and assigning function to the isolated DNA fragment. An additional advantage is the ability to make by in vitro transcription coupled to in vitro translation (There is an ATG in frame with
20 either GAL4DB or GAL4AD) labeled protein which can be used for in vitro controls (e.g. pull down assays) of the actual protein-protein interaction.

The sequences of the plasmids produced and the SP6 and T3 polymerase are identified in the Sequence
25 Listing provided below:

SP6 DNA-dependent RNA polymerase:

SEQUENCE ID NO. 1

swissprot accession number P06221

protein sequence:

```
1  mqdlhaiqlq  leeemfnggi  rrfeadqqrq  iaagsesdta  wnrrllseli  apmaegiqay
61  keeyegkkgr  apralafiqc  venevaayit  mkvvdmlnt  datlqaiams  vaeriedqvr
121 fskleghaak  yfekvkkslk  asrtsyrha  hnvavvaeks  vaekdadfdr  weawpkeqtl
181 qigtllleil  egsvfyngep  vfmramrtg  gktiyyllts  esvgqwisaf  kehvaqlspa
241 yapcviprp  wrtpfnggfh  tekvasrirl  vkgnrehvrk  ltqkqmpkvy  kainalqntq
301 wqinkdvlav  ieevirldlg  ygvpsfkpli  dkenkpanpv  pvefqhlrg  elkemlspeq
361 wqqfinwkge  carlytaetk  rgksaavvr  mvqqarkysa  fesiyfvyam  dsrsrvyvs
421 stlspqsndl  gkallrftg  rpvngealk  wfcinganlw  gwdkktfdvr  vsnvldeefq
481 dmcrdiaadp  ltftqwakad  apyeflawcf  eyaqyldlvd  egradevrth  lpvhqdgscs
541 gqhysamlr  devgakavnl  kpsdapqdiy  gavaqvikk  nalymdadda  ttftsgsvtl
601 sgtelramas  awdsigitrs  ltkkpvmtil  ygstrltcre  svidyivdle  ekeaqkavae
661 grtankvhpf  eddrqdytlp  gaaynymtal  iwpsisevvk  apivamkmir  qlarfaakrn
721 eglmytlptg  fileqkimat  emlrvtclm  gdikmslqve  tdivdeaamm  gaaapnfvhg
781 hdashliltv  celvdkgvts  iavihdsfgt  hadntltlr  alkgqmamy  idgnalqkl
841 eehevrmvd  tgievpeqge  fdlnemidse  yvfa
```

T3 DNA dependent RNA polymerase:

SEQUENCE ID NO. 2

swissprot accession number P07659

protein sequence:

```
1  mniieniekn  dfseielaa  pfntladhyg  salakeqlal  ehesyelger  rflkmlerqa
61  kageiadnaa  akpilatllp  kltrrivewl  eeyaskkgrk  psayaplql  kpeasafitl
121 kvilasltst  nmittiqaag  mlgkaiedea  rfgrirdlea  khfkhhveeq  lnkrhgqvyk
181 kafmqvvead  migrgllgge  awsswdkett  mhvgirliem  liestglvel  qrhnagnags
241 dhealqlaqe  yvdvlakrag  alagispmfq  pcvvppkpww  aitgggywan  grrplalvrt
301 hskkglmrye  dvympevyka  vnlaqntawk  inkkvlaavn  eivnwknpcv  adipslerqe
361 lppkpddidt  neaalkewkk  aaagiyrldk  arvsrrisle  fmleqankfa  skkaiwfpyn
421 mdwrgrvyav  pmfnpqgndm  tkgltlakg  kpigeegfyw  lkihancag  vdkvpfperi
481 afiekhvddi  lacakdpinn  twaeqdspf  cflafcfeya  gvthhglsyn  cslplafdgs
541 csgiqhfsam  lrdevggrav  nllpsetvqd  iygivaqkvn  eilkqdaing  tpnemitvtd
601 kdtgeisekl  klgtstlaqg  wlaygvtrsv  tkrsvmtlay  gskefgfrq  vlddtiqpai
661 dsgkglmftq  pnqaagymak  liwdavsvtv  vaaveamnwl  ksaakllaae  vkdkktkeil
721 rhrcavhwtt  pdgfpvwqey  rkplqkrldm  iflgqfrlp  tintlkdsi  dahkqesgia
781 pnfvhsgdgs  hlrmvvyah  ekygiesfal  ihdsfgtipa  dagklfavr  etmvityenn
841 dvladfysqf  adqlhetql  kmpplpkkgn  lnldilkds  fafa
```

pGN108:

SEQUENCE ID NO. 3

[illegible]

tctctcttgactctctcttccgccaaaagcctagcattttattgataatttgattacacacactcagagttctcgcacatgataaagtgttctatggcac
tcgccctaacagtacatgacaaagggcggaattattatcgaatgataatgaagacaaactccaaatgtgtctcattttggagcccgctgtggggcagctg
ctctcaatatattactaggagacgaggagggggacattatcgaacgtcgaatgagccattcttctctttatgcaactcttcaactctctcacacattaat
cgattcatagactcccatattcttgaatgaaggtgtgggttttagctttttcccgattgttaaaaggaagaggctgacgatgttagaaaaagagaacg
gagccgaaaaaacatccgtagtaagttctcttttaagccgacacatttttagacagcattcgcgcgtagtttgaaagttaaattttaaaaataaaaaatag
ttcaatttttttaattactaaataggcaaaagtttttcaagaactctagaaaaactagcttaattcatgggtactagaaaaattctgttttaatttaatttta
cttaagattgtaaatcagagaagctttttgaaaaaattctcaattaaaaagatttgcggttaggataaaaaagctctcagaaatgagttaaagctcaaatgga
agttgtttttaaaagaaaaaacgaaaaaaagacacactattttcttccctccgcgttaaaattagtgttgtagataatgtgatccgcgtctgtatttgcact
cggctcttcacacgcgtctctctcacttgaccacagtgaaaaaaaacactcagctctgagacgttgaaattgccttattcagcgcgtctgtcttcca
cccagtaacaaaaaaaattggmtcttattttatattatgtagggtcacaaaaaaaagtgatgcagttttgtgggtcggtgtctccacacacacctccgc
ctccagcagcacacaatcattctgtgttctcgcagatctctgtatgccggctcgtgaatgcaccacattcgcgcgcaactcacaccacactc
actttcgggtggtattactacacgtcatcgtgttctcgtagctctccgctcttctcgtcccaactcactctcattattcccttgggtgtattgtttttaaatggta
caccactctcagcttttactcttctgttttccgtccatttagattttactggaatttttttaaaatttttagccagagagttctagtctgttcttaaaagtcta
ggtcagacatacatttctatttctcatcaaaaaaaaagttgataaagaaaactgggtatttcagaaaagagtgctgtcgtggaattgattcaaaaaaaatt
ccacccctcgtgttttttcaaaattgagatcaaacaggatttttcttctcgtatcaatttttctgctgcgtctgtgccaagtgtgtgtgtccgagcaaa
agatgagagaattacaacacagaaatgtaaaaaagcttggccaaataatgaagtgtttatcccgagatgattggaagaatataaattgtctttacgtgttga
ggggagagagagatagattttcgtatcaaaactccgcttttacctatgtcttttagaatcctaaatagatttttctatctttaaattgaaaaactgagaaaat
cagtaatttcgcaatttcttgcataaaatacacgaaattgtgggtctgccacgatctcggctcttagtgggtcatttgggttaaaggtttaaaatttcaaa
ttctagtgttaaatttccgcalaattggacctaaaatagggttttgcatttcaacaagaatactgtaaaactcgtgtgttccgcaatttttcaaaaata
cacgaaatataatggtaatttcccgaaataatgagggtctcgcacagatttcagtcacagtggccagatttatcacgaaaaaagttcgcctagtctcacat
ttccggaaaaaccgaatcaaatagtttttgcattcatattgaaacaaaaatctgagacatccctatagtttcgaattttcgtcgttttcttccaaaaatga
cagcttagaataaaaattcgttgaactgggacctatgatacttttctcccggttttactttttttttttttttacacttggaattgactaaaggtcaccaccaccg
cagttgtgttgccattcacacacacacacacacagatctcgagatttttctgttctccctgtctgatttccgttctctctctctcttattcatctttt
agccgagaagctccagaaatgagacacacacacactccggcgcgagctctgggagagctgcgcgcctgggaagccgcccagagagatatca
gggaagatcgtctgatttctctcggatgccacctatctctcagtttctcgcctgttactccttccggaacctgataatttccc

pGN105:

SEQUENCE ID NO. 4

aagcttgcacgtcctgcaggccttgctgactctagacacttttcagctacctagatagatgataatccccgcctcccaatccaccaccaggggaaaa
gaagggtcgcggcaaaaatcaagttatctccaggctcgcgcacccaccgagcgggtgacttctccaccacttttcatthaacctcgggttacg
ggatggccaaaggaccacaaaggatgtttcgaatgatactaacatacatagaaacatttcaggaggacccttgcttgagggtaccgagctcagaa
aaaatgactgtctcaaaagaagaagcgtatagggtaccggtatgaacacgattaacatcgtctaaagacgacttctctgacatcgaactggctgctatccc
gttaacactctggctgaccattacgggtgagcgttttagctcgcgaacagttggcccttgagcatgagtcctacgagatgggtgaagcacgcttcgcaa
gatgtttgagcgtcaacttaagctgggtgagggtgctgataacgctgcgcgaagcctctcatcactaccctactccctaaagatgttgacgcatcaa
cgactgggttgagggaagtgaagcgttaagcgcggcaagcgcggcagcagccttcagttcctgcaagaaatcaagccggaagcgtgacgtacatca
ccattaagaccactctggcttgccctaaccagtgtgacaatacaaccgttcaggctgtagcaagcgaatcggctgggccattgaggacgaggctcg
cttcggctgataccgtgacctgaagcgaacacttaagaaaaacgttgaggaaacactcaacgaagcgcgtaggggcacgtctacaagaagcattta
tccaagtgtcgcagggtgacatgctcttaagggtctactcgggtggcggagcgtgtgtcttctggtgcataaggaagactctattcatgtaggatgacgt
gcatcgagatgctcattgagtcacccgaatgggttagcttacaccgccaataatgctggcgttagtgtaggtcaagactctgagactatcgaactcgacct
gaatcgtgagggtatcgcaacccgtgagggtgctggctgcatctctcgatgttccaaccttgcgttagtcttccctaaagccgttggaactggcatt
actgtgtgtgtgctattgggctaaccggtgctgctcttggcgtggtgctgactcacagtaagaagacactgagctgctacgaagacgtttacatgcct
gagggtgtcaaaagcattaacattgcgcaaacaccgcatggaaaatcaacaaagtcctagcggctgccaacgtaacccaagtggaagcat
tgtccggtcgaaggacatccctgcgttggagcgtgaagaactcccgatgaacccggaagacatcgacatgaatctgaggctctaccgcgtggaaa
cgtgctgcccgtgctgtgtacccgaaggacagggctgcgaagtctgcgcgttatcagccttgagttcatgcttgagcaagccaataagttgttaacat
aaggccactctggttcccttaacaatggactggcggcgtgtgtttacgcggtgtcaatgtcaacccgcaaggtaacgatagcaaaagactgctt
acgctggcgaaaggtaaaccaatcgttaagggaaggttacttactggtgaaatccacggtgcaaacgttgcgggtgtcgaataaggttccgttccctg
agcgcatcaagttcattgaggaaaaccacgagaacatcatggcttgcgttaagtctccactggagaacacttgggtgggtgagcaagattctccgttct
gttccctgcttctgctttagtactgctgggtgacagcaccgcccgtgagctataactgctcccttccgctgctggcgtttgacgggtctgtcttggcact
cagcacttctccgcatgctcggagatgaggtaggtgggtgcgcgggttaacttcttctagtgagaccgttcaggacatctacgggattgtgtgctaag
aaagtcaacgagatttacaagcagacgcaatcaatgggaccgatacgaagtagttaccgtgaccgatgagaacactggtgaatctctgagaaa
gtcaagctgggcaataaggcactggctgtaaggtgctgacgggtgactcgcagtgtagtaagcgttcagtcacgtgcttaccgggtc
caaagagttcggcttccgtcaacaagtgtcgaagatacattcagccagctattgattccggcaagggtccgatgttcactcagccgaatcaggctg
ctggatactggctaagctgattgggaatctgtgagcgtgacgggtgtagctgcgggtgaagcaatgaactggttaagtcgtgctaagctgctgg
ctgctgagggtcaaaagataagaagactggagagattcttgcgaagcgttgcgtgctgattgggttaactctgattggttccctgtgtggcaggaataca
agaagccctattcagacgcgcttgaacctgatttctcgtcagttccgcttacagccctaccattaacaccaaaaagatagcgagattgatcacaca
aacaggagctgtgtatcgtccttaactttgtacacagccaagcgggtgtagccaccttcaagactgtagtgtggcagcacgagaagtacggaatcga
atctttgactgattcagactccttccgttaccattccggctgacgctgcgaacctgttcaaaagcagtgccgcaaacattgtgtgacacatagagctt
gtgatgtactggctgatttctacgaccagttcgtgaccagttgcacgagtcataatggacaaaatgccagcacttccggctaaaggtaacttgaacct
ccgtgacatcttagagtcggacttccgttgcgttaagaattccaactgagcgcgggtgctaccattaccaacttctgtgtgtaaaaaataatagg
ggcgtgtcatcagagtaagttaaacgtgatttctactaactaacgagtaataatttaatttcagcatctcgcgccgtgcttctgacttctaagtcacat
acttccaacatccctacatgctcttctccctgtgtctccacccttattttgtattatcaaaaaaaccttcttatttcttgtttttgacttcttaagtc
ccttaacaatgaaattgtgtgattcaaaaatagaatattcgttaataaaagtcgaaaaaattgtgtccttcccccttaataataattctatccca
aaaactcaacaatgttctgtgtacacttcttattgttttttacttctgataattttttgaaacatcatagaaaaaccgcacaaaaatccttatcatgtt
acgtttcagtttatgaccgcaatttttatttctgcacgcttggcccttctacagtcacaaatcatgctcactgtgaaaaagtttggagatttttgaatttt
caatcaagtgaagttatgaaatatttctgcttttgggttcccttattgttttcaagagtttgcaggacggcgttttcttctgtaaaatca
caagtattgatgacgacgatgcaagaaagatcggaagaaggttgggttggagctcagtggaaggtgagtagaaggtgataatttgaagtgaggta
gtgtctatgggggttttccctaaatgacagaatacattcccaataacacataactgttcttactagtcggccgtacggcccttctgctcgcgcg
ttcgggtgatgacggtgaaaacctctgacacatgcagctccggagacgggtcacagcttctgttaagcggatgcggggagcagacaagcccgta
ggggcgctcagcgggtgttggcggtgtcggggctggcctaactatgcggcatcagagcagattgtactgagagtcaccatatgcggtgtgaaata
ccgcacagatgcgtaaggagaaaataccgcatcagcggcccttaaggccctgtgatacgcctattttatagggttaattgtcatgataaataatggtttct
agacgtcaggtggcacttttggggaaatgtgcgcgggaacccctatttatttttctaaatattcaataatgatatccgtcatgagacaataaccct
gataaatgttcaataataattgaaaaaggaagatgatgatttcaacattccggtgtcgccttattcccttttggcgacttttcttctgttttctca
cccagaaacgtgtgtgaaagtataaagatgctgaagatcagttgggtgcacgagtggtttacatcgaactggtatcgaacagcggtaagatccttgag
agtttctccccgaagaacgttttccaatgatgacacttttaagttctgtatgtgtgcgcggtattatcccgatttgcagccgggcaagagcaactgc
gtcgcgcatacactattctcagaatgactgtgtgacttaccagtcacagaaaagcacttaccgatggcatgacagtaagagaattatgacgtg
ctgccataacctgagtgataacactgcggccaacttcttgcacaacgatcgaggaccgaaggagtaaccgctttttgcacaacatgggggat
catgttaactcgccttgcgtgtgggaacgggagtgatgaagccatacacaacgacgagcgtgacaccacgatcctgtagcaatggcaacaacg
ttgcgcaactattaactggcgaaacttacttactgttccggcaacaataatagactggatggaggcggaataagttgcaggaccacttctgcgt
cgcccttccggctggtgtttatgtgataaatctggagccggtgagcgtgggtctcgcgttatcattgcagcactggggccagatggttaagccc
tccgtatctgattatctacacgacggggagtcaggcaactatggtgaacgaaatagacagatcgctgagataggtgctcactgattaagcattg
gttaactgtcagaccaaagtttactatatacttttagattttaaacttattttaaagagatctaggtgaagatccttttataatctcatgacca
aaatcccttaacgtgagtttcttccactgagcgtcagaccggtagaaaagatcaaggatcttcttgagatccttttttctgcgctaactctgctgt
gcaacaaaaaaaccaccgctaccagcgggtgtgtttgttgcggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgcaga
taccaaaatctgtcttctagtgtagccgttagtgccaccacttcaagaactctgttagcaccgctacatacctcgtctgttaactctgtaccagtg
gctgtgtccagtggtgataagctgtgttaccgggtggactcaagacgatagttaccggataaggcgcagcggctgggtgaacgggggttgcg

tgacacagcccagctggagcgaacgacctacccgaactgagatacctacagcgtgagcattgagaaagcgccacgctcccgaaggagaa
aggcggacaggtatccggtacggcagggtcggaaacaggagcgcacgaggagcttcagggggaaacgcttggtatcttatagtcctgctc
gggttcgccacctctgacttgagcgtcgatcttgatgctcgtcagggggcgaggcctatggaaaaacgccagcaacgggccttttacggttc
ctggcctttgctggcctttgctcacatgttcttctcgttatccctgattctgttgataaccgtattaccgcctttgagtgcgtgataccgctcgcg
cagccgaacgaccgagcgcagcagtcagtgagcggaggaagcggaagagcgcccaatacgcaaaccgctctccccgcgcgttgccgattca
ttaatgcagctggcagcacagggttccgactggaaagcgggcagtgagcgcgaacgcaattaatgtgagttagctcactcattaggcaccagcgtt
tacactttatgcttccggctcgtatgtgtgtggaattgtgagcggataacaatttcacacaggaaacagctatgacctgattacgccaagctgtaagtt
aaacatgatcttactaactaactatttcatttaattttcagagcttaaaaatggctgaaatcactcacaacgatggatacgttaacaacttggaatgaa
at

pGN400:

SEQUENCE ID NO. 5

aagcttgcatgacctgagggccttgctgactctagacacttttcagctacatgatacatggatatccccgctcccaatccacccaccaggga
gaagggctcgcgcaaaaaataaggtatctccaggctcgcgcacccaccgagcgggtgacttctccaccacttttcatttaacctcggggtag
ggattggccaaaggacccaaggtagtttcgaatgatactaacataagaaacatttcaggaggaccttgctggagggtaccgagctccgg
gattaatcagctactataccgtagaaaaatgagtaaggagaagaacttttcactggagtggtcccaattctgtgaattagatggtgatgtaag
ggcacaattttctgtagtgagggtgaagggtgatgcaacatcggaacacttacccttaatttttgcactactggaaaactacctgttcattgg
gtaagtttaacatataatactaaccttgattttaattttcagcccaacactgtcactactttctgttaggtgttcattctcgagataccag
atcatatgaacggcatgacttttcaagagtgcacgtcccgagggtatgtacaggaagaactatattttcaagatgacgggaactacaagcac
gtaagtttaaacagttcggtactaactaacatataatttaatttcagggtgtgaagtcagggtgataccctgttaataagaatcgagttaaa
agggtattgatttaagaagatggaaacattcttgacacaaaattggaatacaactatactcacacaatgtatacatcatggcagacaaaacaaagaat
ggaatcaaaagtgttaagtttaacatgattttactaactaactaatctgatttaattttcagaacttcaaaattagacacaacatgaagatggaggttca
actagcagaccattatacaaaaaactccaattggcgatggccctgtccctttaccagacaacattacctgtccacacaatctgccccttcaagatc
ccaacgaaaagagagaccacatggctctcttgagttgttaacagctgctgggattacacatggcatggatgaactatacaaatagcattctgtagaattc
caactgagcgccggctgctaccattaccaactgtctgtgtcaaaaataataggggccgctgtcatcagagtaagtttaaacgtgcttactaactaa
cgagtaatttttaattttcagctctcgcgcctgtccttgacttcttaagtgccaattacttcaacatccctacatgctcttctccctgtctccacc
cctattttgtattatcaaaaaacttcttcttaattttctgttttttagcttcttaagtcacctctaacaaatgaaattgtgtagattcaaaaatagaattatctg
taataaaaagtcgaaaaaatgtgtctccctcccccttaataataattctatcccaaatctacacaatgttctgtgtacacttcttattgttttttactctg
ataaatttttttgaacatcatagaaaaaccgcacacaaaatccctatcatatgttacgtttcagtttatgaccgcaatttttttcttctgcacgtctggc
ctctcatgacgtcaaatcatgctctatcgtgaaaaagttttgagtagttttgaaattttcaatcaagtgaaagtattgaaatttttctgtctttttg
gggggttccctattgtttgcaagagtttcgaggacggcgtttttctgtcaaaatcaagaattgtatgacacgtgcaagaagatcggaagaagg
ttgggttgaggctcagtggaagggtgagtagaagtgataattgaaaggtagtgatgtctatgggttttgccttaaatgacagaatacattccaat
ataccaaacataactgttctactagtcggccgtacggccctttctctcgcgcgtttcgggtgtagcgggtgaaaccttgacacatgcagctccc
gagacgggtcacagcttctgttaagcggatgcgggagcagacaagcccgtagggcgctgtagcgggtgttgaggagaataaccgcacatgcagcggc
actatgcgcatcagagcagattgtactgagagtgacacatgctggtgtgaaataccgcacagatgcgtaaggagaataaccgcacatgcagcggc
cttaaggccctctgatacgcctattttataggttaattgtcatgataataatggtttcttagcgtcaggtggcacttttggggaaatgtgcggaacc
cctattgtttttttaaatacattcaaatatgtatccgtcatgagacaataaccctgataaatgtctcaataatattgaaaaaggagatgagatt
caacatttccgtgtcgccttatttcccttttttgcggcattttgccttctgtttttgtcaccagaaacgtggtgaaagtaaaagatgctgaagatcagtt
gggtgcacgagtggttacatcgaaactggatctcaacagcggtgaagatccttgagagtttgcgccgaagacgttttccaatgatgagcattttaa
gttctgtatgtggcgcggtattatccgtattgacggcggaaggaacacgtgctgcgcgcatactattctcagaatgactgtggtgagtactac
cagtcacagaaaagcatcttacggatggcatgacagtagaagaattatgcagtgctgccaataaccatgagtgataacactgcggccaacttactctga
caacatcgaggagcgaaggagctaacgcgtttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagc
cataccaaacgacgagcgtgacaccagatgcctgtagcaatggcaaacgttgcgcaaaactataactggcgaactacttactctgacttccggc
aacaattaatagactggatggaggcggaataagttgcaggaccacttctgcgtcggccctccggcgtggtttattgtgataaatctggagccg
gtgagcgtgggtctcgcggtatctgacgactggggccagatggttaagccctccgtagttagttatctacacgacggggagtcaggcaactat
ggatgaacgaaatagacagatcgtgagataggtgctcactgattaagcattggttaactgtcagaccaagtttactcataatacttttagattgatttaa
acttcattttaatttaaaaggatctaggtgaagatccttttgataatctcatgacaaaatcccttaacgtgagtttctgtccactgagcgtcagaccccg
tagaaaagatcaaaaggatcttcttgagatccttttttgcgcgtaatctgtctgttgcacaaaacacccaccgctaccagcggtgtgttgcgg
atcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgagatacaaaatctgtcttcttagtgtagccgtagttaggccaccact
caagaactctgtagcaccgctacatacctcgtctgttaacctgttaccagtggtctgctgcgataagtcgtgttaccgggttgactca
agacgatagtaccggataaggcgcagcgtcgggtgaacgggggttcgtgcacacagccagcttgagcgaacgacacacccgaactga
gatacctacagcgtgagcattgagaagcgccacgcttcccgaaggagaaaggcgacaggtatccggtaagcgccagggctggaacaggag
agcgacgaggggagcttccaggggaaacgcctggtatctttagctctggtttcggccctctgactgtgagcgtcgtattttgtgctcgtc
agggggggcggagcctatgaaaaacgcagcaacgcggcctttttagcgttctggttctgtgcttctgtcactgttcttctgcttacc
ctgattctgtgataaccgttattaccgctttgagtgagctgataccgctcgcgcagccgaacgaccgagcgcagcagtgagtgagcgaag
cggaagagcgcccaatacgcacaccccttccccgcggttgccgattcattaatgagctgacgacaggttcccgactggaaagcgggc
agtgcgcgaacgcaattaatgtgagttagctcactcattaggcaccacaggctttacattatgttccggctcgtatgtgtgtggaattgagcgg
atacaatttcacacaggaacagctatgacatgattacgccaagctgtaagtttaacatgacttactaactatttctttaaattttcagagctt
aaaaatggctgaaactcactacacagatggatagcctaacaacttggaaatgaaat

SEQUENCE ID NO. 6

gacacggcgccgcatgctgcgggagatggcgccgcttgggaagcgccgagagatacagggaagatcgtctgatttctccggtgacccatt
catctctcagtttccgcttactccttccgaacctgataattccgtgttcglaaagagatgttttattttactttacacgggtctctctctcgtcc
agcacagctcagtgttggctgtgtcgtggctcctgccacggcgccctcacttcttcttctctctcgtctcgtctatcactttcttattctt
attccttttcatcatcaaaactagcattcttactttatttttttcaattttcaattttcagataaaaacaaactacttgggtttacagcgcaacagatccccg
ggattggccaaaggaccacaaaggatgtttcgaatgatacataacatagaaacattttcaggaggacccttgcgttggagggtaccggtagaaaaa
atgagtaaaaggagaagaacttttactgaggtgttcccaattctgttgaattagatgggtgatgttaattgggcacaaattttctcagtgaggagggtga
agggttgacacatcggaaaacttccgttaatttttgcactactggcaaaactacctgttccatgggtlaaggtttaaacatatatactaactaacgtg
gattttaaatttttccgaacactgttcacttttctgtttatgtgttcaattcttctcagatataacacagcagctagctatttttcaaggagt
gccatgcccgagggttatgtacaggaagaacttatattttcaaaagtacgggaactcaagacacgtgaattttaaaccgttgcgtactaactaacca
acatatttaaaatttcagggtgtcgtgaagtcaagttgaagggtgataccctgttgaatagaatcgagttaaagggtattgattttaaaggagtggaaacattctt
ggacacaaatttgaatacactataactcacacaatgtatacatcatgtgcagacaaaacaaaggatggaatcaagggtgaagttaaacttggactttac
taactaacggattattttaaaatttcagaacttcaaaaattagacacaaatggaagtggaaagcgttcaactagcagaccattatcaacaaaatactccaa
ttggcgatggccctgtcctttaccagacaacacattacctgtccacacaatctgcccttccgaagatcccaacgaaaagagagaccacatgtgtcctct
tgagttgttaacagctgtcgtggattacacatggcatgtgatacatacaatgacattcgtagaattccaactgagcgccgtcgtctaccattaccaac
tgtctgtgtgtcaaaaataataggggccgctgtcatcagagtaaggtttaaacgtgattctactaactaacgagtaattttaaaattttcagcatctcgtcgccc
gtcctctgtacttcaatgtcaacttctcaaacactccctactgtccttctcccttgcgttccacccctatttttggattatcaaaaaatttcttctaatt
tctttgtttttgtccttttaagtacacctcaacaatgaaattgtgtgtattcaaaaattgtcaataattcgttaataaaagctgaaaaaattgtgtctcctcc
ccccattataataattctatcccaaatctacacaaatgttctgtgtacacttctatgtttttttacttctgataatttttttgaacatcatagaaaaaacccg
cacacaaaatacctttatcataatgtacgtttcaggtttatgaccgcaatttttttcttcgcacgtctggccctctcatgacgtcacaatcatgtctatcgtgaaa
aagtttttgagatattttggaattttcaatcaagtgaaaagtattgaaattaaatttctcgttttgcgtttttgggggtttccctattgtttgtcaagagtttcgag
gacggcggtttttctgtcaaaatcacaaagtattgatgagcacgatgcaagaagatcggaagaagggtttgggttgaggctcagtggaagggtgagtag
aagttgataattgaaagtggagtagtgtctatgggtttttgccttaaatgacagaatacattcccaatataccaaacataactgtttcctactagtctggcc
gtacggcccggtacccagcttttgccttttagtgagggttaattgctgcgttggcgttaatcatgtgicacatagctgtttcctgtgaaattgttatccgct
cacaatttccacacaacatcagcggcggaagcgttaaaagccttggcgttccctaattgagtgagcgttaactacataattgttcgttgcctcactgc
ccgtctccagcttgggaacctgtctgtcggcgtgtcattaatgaatcgccgaacgcgggggagaggcgttggcgttgcgttattggcgctctccgcttcc
tgcctcactgactcgtcgtcgtcgttgcgttgcgtcgtcgtcgtcgtatcagctcacaaggcgttaatacgggtattccacagaatcagggggata
acgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtlaaaaggccggtgtcgtcgtcgttttccataggctccgccccct
gacgagcatcaaaaaatcgacgtcaagtcagagggtggcgaaacccgacaggactataaagataccaggcggtttcccttggaaagctccctcgtg
cgctctcgttctccgacctgccgttaccggatacctgtccgcttctccttccgggaagcgttggcgttttctcatagctcacgtctgtaggtatctcag
ttcggtgtaggtcgttgcctcaagctggcgtgtgtgcaggaacccccgttcagcccagccgtcgtccttatccggtaactatcgtcttgagtccaa
ccccglaagacacagcttatcggcactggcagcagccactggttaacaggattagcagagcgaggtatgtaggcggtgtctacagagttctgaagt
gtggcctaactacggctacatagaaggacagctatttggatctcgtcgtcgttgaaggcgttaccttccgaaaaaaggttggtagctcttgatccgg
caaacacacaccgctgtgtagcgggtgtttttgttttgcgaagcagagatnctgcgaagaaaaaaggtatcgaagaagtatttggattttttacg
ggctcgtacgcctagtgtagcgaagaaactcagttaaagggtttgtctatgagattcaaaaaaggtatccactagatctttaaataaaaaaggaag
tttaaatcaatcaaaagtatatatgagtaaactgtgtctgacagttaccaatgcttaatcagtgaggcgacctatctcagcagatgtctatttcttcatccat
agttgctgactccccgtcgtgtagataactacgatacggagggtttaccattcgtcccgagtgtcgtcaatgataccgcgagaccacgctcaccg
gtcaccagattatcagcaataaccagccagccggaaggggcgagcgcaaggtgtcgtcgaactttatccgcttccatccagcttattaattgttg
cggaagctagagtaagtgtcgcagtttaattgttgcgaacgttgggtgcaattgtctacggcagctgtgtgtgtcagcgtcgtcgtgttggtatggctt
attcagctcgggttcccaacgatcaaggcgagttacatgatccccatgtgtgtcaaaaaaggcgttagctccttgcgttccgctcgtgtgtcagaagt
aagtgtggcgcagtggtatcactcatgtgtatggcagcactgcataattcttacttgcattccatccgtaagatgcttttctgtactgttgtagtactcaa
cctatcattctgagaatagtgtatgcggcgaccgagtgtcgttcttccggcgctcaatcagggtatataccgccaacatgacagaactttaaagt
ctcaatcgttggaaaactgttctcgtggcgcaaaactcgaagctatccactgtgttgagatcggcttcatgtgaacccactgacacccaactcactt
agcatctttactttaccagcgttttgcgttgtagcaaaaaaggaaaggcaaaatgcgcaaaaaagggaataaggcgacacgggaagtgtgtaata
ctcatacttcttcttttcaatattatgaagcatttatcagggttatgttctcatgagcggatacatatttgaaatgtatttagaaaaataacaaatagggggtc
cgcgacatttccccgaaaagtgcacctaattgtaaagcgttaataattttgtaaaatcgcgttaaaattttgttaaatacgtcatttttaacaaataggc
cgaatcggcaaaatcccttataaatcaaaagatatagaccgagatagggtgtgtgtgttccagtgttgaacaagagctccattataaagaacgttga
ctcaacgtcaaaaggcgcaaaacgtctatcaggcgatgtgccactacgtgaaccatcaccctaatcaagtttttgggtgcagggtgcggttaaa
gcactaaatcgaacccataaagggaagccccgatttagagctgtacggggaagccggcgcaacgtggcgagaagggaagggaagaaagcgaa
aggagcgggctgtagggcgctggaagtgtagcgggtcagctgtcgtgaaccaccacaccccgccgcttaattgctgcgctacaggggcgctc
ccattcgaacttaccgttgcgaactgttggcggaaggcgatcgtgtgcggccttctcgtattacgccagctggcgaaaagggggagatgtctgcaag
cgattagggttgggttaacgcgaattgttttccagctacagcagttgtaaaacgacggccagtgagcgcgctaatacgaactcactatagggcgaaatt
gagctccaccgggtgtgcggccgctctagaactagt

pGN110:

SEQUENCE ID NO. 7

gatctccaaaatgcttccgctctgaaaaacgaagtgacacccgaaataatggcgcaaaaatgaaatgagcttttgggtcgaaaa
aaatgttttagaatgctgagaaacgtaaacacgaagatcatatttttgagaccggatgctctgaaaaatgctgacatagattaaaaagcatat
atatatttttcatttcaacgtgaaagtttgcaactttatagaatctctatttgacacattgtttttttaactgaggcagttttgaaacaccttttgaaact
tgaatctcttgaaagtatactgctgaaaaagactgactgagcggttgaaatgccagaagaaaactataattgaaatcgcgcgtaaattgagaaatgcaacc
gctgctccactggacaattgaaaaaaatttattcggagggcacaacggatatttcgaaatgattttctgtatttttcaattttataaattcttcttgatt
atcgcttctgttgagaaatttaattgattcaaaactttttatagtaagataccgggtgtaccgtagcgtacgaacccgggattggccaaaggacca
aaggatgtttcgaatgataactaacaatagaacattttcaggaggacccctgtgtggagggtaccggatgactgctcaaaagaagaagcgtaagc
tcatgaacacgattaaatcgtctagaacgacttctgacatcgaactggctgtatccgttcaacacactgtgctgaccattacggtgagcggttagct
cgcgaaacagttggcccttgagcatgactttagagatgggtgaagcagcgttcgcaagatgtttgagcgtaactaaagctgggtgaggttgcgga
taacgtgctggcccaagcctctcatcactaccctactccctaagatgattgacacgcatcaacgactggtttgaggaagtgaagcctaagcgcggcaag
cgcccagacgcttccagttctgtcaagaaatcaagccggaagccgtagcgtacatcaccattaaagaccacttggcttgcctaaccagtgctgaca
atacaaccgttcaggctgtgcaagcgcgaatcggtggccattgaggacgaggctcgttggctgtatccgtgacctgaagcctaagcacttcaa
gaaaaacgttgaggaaacactcaacagcgcgtgaggcacgtctacaagaaagcattatgcaagtgtcagggtgacatgctcttaagggtctta
ctcggtggcgaggcggtgttcttggcataaggaaagacttattcatgtaggagtagcgtgacatgagatgctcattgagtaaccggaaatggttag
cttacaccgccaatgctgctgtaggtgcaagactctgagactatcgaactcgcaccgtgaatcgtgaggctatcgaacccgtgacgggtgcg
ctggctggcatctctccgatgttccaacttgcgtagtcttcttaagccgtgactggcattactggtgtgtgttattgggtcaacggtcgtctct
ggcgctgtgtgctactcagtaagaaagcactgatgctcagaaagcgtttacatgcttgagggtgtacaagcgattaaacattgcgcaaacacc
gcatgggaaaaatcaacaagaagctctagcgggtgcgaacgtaatacacaagtggaagcattgtccggtcaggacatccctgcgaatgagcgtgaa
gaactcccgatgaacccggaagacatcgacatgaatcctgaggtcctcaccgctggaacgtgtgctgcccgtgtgtaccgcaagacaaggtc
gcaagctcgcggtatcagccttgagttcatgcttgagcaagccaataagttgttaaccataagccatctgttcccttaacatggaactgcggtc
gttcgtgttacgctgtgcaatgttcaacccgcaaggtaacgatatgacaaaggacgtcttacgctggcgaaggtaaaccaatcggtgaaggaag
ttactactggctgaaatccacgggtgcaaatgtgcgggtgtcgataaggtttcttctgagcgcacaaagttcattgagaaaccacgagaacat
catggcttgcgtaagcttccactggagaacacttgggtgagcaagattctcgttcttctgcttctgttctgattgacgctgggtacagc
accacggcctgagctataactgtctcccttccgtggtgtttagcgggtctgtctgcatccagcacttctccgcgtatccgcagatgaggtagggtg
gtcgcgcggttaactgttcttctagtgaaacccgttcagacatctacgggattgtgttaagaaagcaacgagattctgcaagcagacgcaatcaatg
ggaccgataacgaagtagttaccgtgaccgatgagaacactgttgaaatctcigagaaggtcaagctgggcactaaggcactggtgtcaatggct
ggcttaccggtgttactcgcaggtgtactaagcgttcagctacgctggttctacgggttccaaagagttcggcttccgtaacaaggtgtggaagatac
cattcagccagctattgattccggcaaggggtctgatgttctacagcgaatcagcgtctggtgatacattggttaagcgtgaggtgagagattctt
gacgggtgtgtagctgcgggtgaagcaatgaactggcttaagctgtctgtaagctgctggtgaggtcaagataagaagactggagagattctt
gcaagcgttgcgtgtgattgggttaactctgatgttccctgtgtggcaggaatacaagaagcctattcagacgctgtgaacctgatttctccgg
tcagttccgcttaccagcctaccataacaccaacaagatagcgagattgatgcacacaacaggagctgtgtatcgtcttaactttgtacacagcca
agacggtagccaccttctgaagactgtgtgtggcacacgagaagtagcgaatcgaatctttgactgattcagcactccttccgtaccattccggc
tgacgttgcgaacctgttcaaaagcagtgccgaacataggttgacacatatgattctgtgtgattgctggtgatttctacgaccagttcgtgaccag
ttgcagagcttcaattggacaaaatgccagcacttccggctaaaggtaactgaacctccgtgacatcttagatcggacttcggttcgcgttaaggg
ccctcgtgagctgggtcacaatcactgaaactccaaagcagccagctgaggaacgtgaagaagaagaaagagtcattgacaggtttgatttt
cttttggctcaaaaagatgaaattattgattttcagcagatctcccaaaactagcagcgagaagctgtgcaagctgttcacagctgcccaaggaatcgc
gggaagtgtgccaagaggtatgttttcaaaaatcaataactgataataattttttgtttgtgaatttaagaaaataatttcgaaaattccttgaattat
caagattgcagtaatttcgagaaaaattgagataattcatagagctattgttaaatttttgatttcagactgaaacttcggaatacaagagaaatcaa
agaaaaggtgacggggatgatcagcctggcacaccgaacagctatagaagccgggaaacttcaccagctccaaaaagggtccaaggagaccag
gtttgtcaaaagcttctgcgattatttcatttcaattttcagagaatcagagctctcgtgaaaaatcccgggttcgttaagatctccagaaggtcttca
gcacgttccccgtcagatctctagacggcgccgagaagaagctcagaagaagcaatccgaagaccagcaccgtaccagagaaaaaga
agaaaagccgctggtatttctacgaacaagaacggaggagcataattcccccgaacacttcgacttatgcaacaacagattagtgataagca
aagtgaacagatcagagaatgaattgggaagaatgaagaaaaagattacggattgttaacagagtcacgcgaagaatctgttcaaatgtca
gagaacttctcaagagaatgtgattcgttcaaggtgagtgagaaaatcgaaggaaaaagaaagaaatgaattttttcaggggacttctctgcccgtg
acattattcaagctcaggcttctaccaggatttctaacgtctatgacgcttggcggcagttatcaactcgaattccctcatgtcgttgaacttcttctc
cgtcgtctgattgtacagttcaaaagatttccgtagaatgacagaggcgtcaggtgaacgtgatcaaatcatcgcacattgataatcaacaag
ttgctcacgaagtcttgcgtggaatcatgattctgatgttgaagaaccaactgatattcagttgaagtgcgcatgttcttctgaaagagtggtgga
gcaaaagcttctggagattgtcagcagcttcaaacgtgtctacaccgttcttctgcaattctcatggaactgaaagatcggaatgcaactggaatc
gacgtattcagttatgattgagactgcaatcgaattcgaaggaacaaatttgcgtaaggtagaataataatagttattagaaaaataaattag
aataatttaattctactagccaatcaggcgaccttttgcgcatagttctatttgaataatttgagaaatttctatttctcgtcggaaatctggaatt
cgacgagatcttctggttctgtcagctgcatcgttctgtccttcttctgctgttctgtgtacaccaagaacctgttgattcatcaactgaatct
gtgactgtgctgttctcactggatgactagcagcactgattctcgagaaatcagattgagttgcgattagggtgacctagaattgggaataatacgaa
ctttgaaaaatattcaggaggaatgaaaaaatttctcgacaatcctcaaaatttattgaccattgttctcccaacttttcaaaaagttatgaaaa
aatgtgaaaaatcggaattggcaattttcagaccatttttaagcattttcaaaaaaaattgcagctgaaataaatgtctattttcagataaatcgcagcgatt
ttctgtgtctgacactagttttatgaaaaaattgttggaagaacatgtgtgcaataggttaattcagataatttccatgtgtttttttcaattaaccaattatc
caaatcttcaaaactcacaatttgcggagctgggctatcaagaatctgtcgtcagtttataagacgagcatctgtatcactgaaaaattattttaataca
aaactgaaatcaacaaaccacttataactttctgacttctgtctgttcgtgtacgatgacgggtgaagaagccaattgtagtagttgattgttcaagt

[illegible]

cagcttgagcgaacgacctacccgaactgagatacctacgcgtgagcattgagaaagcgccacgcttcccgaaggagaaaggcggacagg
tatccggtaagcggcaggggtcggaaacaggagagcgcacgaggagcctccaggggggaacgccctggtatctttatagtcctgtcgggttcgccac
ctctgacttgagcgtcgattttgtgatgctcgcagggggccgagcctatggaaaaacccagcaacgcggccttttacgggttcctggcctttgct
ggcctttgctcacatgttcttctcggttatccctgattctgttgataaccgtattaccgcctttgagtgagctgataccgctcggcagccgaacga
ccgagcgcagcagtcagtgagcgggaagcgggaagagcgccaatacgaaccgcctctcccgcgcgttggccgattcattaatgcagctg
gcacgacaggttcccgaactggaaagcgggcagtgagcgcaacgaattaatgtgagttacctcactcattaggcaccgccaggttacactttatgct
tccggctcctatgtgtgtgaattgtgagcggataacaattcacacaggaaacagctatgacctgattacgccaagctcgggaattaaccctcactaa
agggaacaaaagctgggggg

gacccgacagagatccctatagtgagctgtattactgcagccgaagcctaattccggcgcaattcttatgattatgattattatataaataagttataaaaa
ataagtgtatacaaaatttaaaagtgcctcrtaggttttaaacgaaaattctgtcttgagtaaccttctgttaggtcagggttcctcagggtatagcatgagggt
cgctcttattgaccacacctctaccggcatgcaagcttggcgtaatactgtgtcatagctgttccctgtgtgaaattgtatccgctcacaattccacacaacatc
gagccgggaagcataaagtataaagccctgggggtccctaataagagtgaggtaactcacaattaattcggttgcgtcactgcccgctttccagctgggaaacctgt
ctggcgctctggaataatgatactggccacgctcggggagaggcggtttgcgatttggcgctctccgcttccctcactcagctacgtcgtcgtcgtcgtcgt
tgcgtcgtcggcgagcgatgcagctcactcaaaagggtglaatacgtgttaccagaatcagggtacacgcaggaagcaatgtgagcagaagggcc
agcaaaaggccggaacgtaaaaaggccggtgtcgtggctgtttccataggctccgccccctgacgagcatcaaaaaatcgacgtcgaagtcagag
gtggcgaaacccgacagagactataaagataccagggtttccccctggaaagctcccctgctgcgtctcctgttccgacctgcccgttaccggataccgttc
cgcttctccctcgtgggaagcgttggcgctttctcatagctacgctgttaggtatctcagttcgggtgtaggtcgttcgtccaaagctgggctgtgtgtcacgaac
ccccggttcagcccagccgtcgcgcttatccggtaactatcgtcttgagctcaaccggtaagacacgacttatccgccatcggcagcagccactgtaaca
ggattagcagagcggaggtatgaggcgtgtcclacagagttcttgaagtgtgtggccttaactcagctacacgtgaagggacagatttggatctgtcgtcgtct
gagccagattccttcgcaaaaaggagttgtgtatcgttctgcggcaaaccaaccacggctgtgtagcgttggttttgttgaacgacagcagattacgcgc
agaaaaaaaggatctcaagaagatccttgtatctttctacggggctgacgctcagtggaacgaaaaatcacgttaagggttttggctatgagattacaaa
aaggatctcacctagatccttttaataaaaaatgaagttttaatacaatcaaaagtatatatagtaaacctgtgtctgacagttaccaatgtcttaacagttaggc
acctatctcagcgatctgtctattcgttaccatagatgtgcctgactccccgtcgtgtagataactacgatacgggaggggcttaccatcttggccccagtgctgc
aatgatacccgagaccacacgctcaccgggtccagattatcagcaataaacacgagccgggaaggccgagcgcagaaggtgtctcgtcaactttatcc
gctccatcacgctattaattgttcgggaagctagatgaagtagtctgcaggtatattgttcgcaacgtgtgttgcattcgtacagagcttctgtgtgtc
cgctcgtctgttggatgtgtcttaccatgtgttcccaacatcaagcgcgagttacatgtccccctgtgttgcataaaaaagggtatcgtctcgtcgtct
ccgatcgtgtcagaagtaagtggccgagtgattatcactatggttatgtgcagcactgcataattcttactgtctcatccatccgtaagatgtctttctgtac
tggtagtactacaacagtcattctgaagaatagtgtatgcggcgaccgaggtgtccttgcggcgctcaatacgggataataccgcccacatagcagaac
ttaaaagtgtctcatcattggaacagcttctcggggcgaaaaactcgaagatcttaccgctgtgtgagatccagttcgatgtaaacctcgtgcaccacaactg
atctcagcatcttttttaccatcagcgttctgggtgagcaaaaacgaaaggcaaaatgccgcaaaaaagggaataaaggcgacagcggaagtgtgaat
actacatcactctctttcttctacattagcagcttattgtgtatgtcttgcagcagatacattggaattgatttgaataaaaaataaacaaatgggggttcgc
gcacattccccgaaaaagtgcacctgaacgaagcatctgtcttattgtgacaaaaaatgtcaacgcgagcgtcaatttttcaaaaaagaatactga
cgtcatmtttagacagacgaatgcaacgcgaagcgtctattttaccaacgaagaatctgtgtcttatttttgaacaaaaaatgacagcgagagcgctaa
ttttcaaacaaagaatctgagctgcatttttacagaacagaatgccaacgcgagagcgctattttaccaacaagaatctatactctttttgttctacaaaaatg
catcccgagagcgctatttttctacaaagcatctttagattactttttctccttgtgcgtctataatgacgtctctgtgataacttttgcactgttaggtccgttaag
gttagaagaaggctcattttgtgtcttcttcttctccataaaaaaacgctgactccacttcccgctttactgtattactagcagaagctgcccgggtcattttcaag
ataaaggctccccgattatttctatccttagctgtgattgtgcatacttctgtgaacagaagaatgatagcgtgtgatgtacttctatgttgcagaataatgaacg
gttctcttattttgtctctataactacgtataggaagtgttaccatttctgtatttctgttctgacttctactacgtatgttgcataaagagtaataac
tagagataaacataaaaaatgtlagaggctgagtttagtgcaagttcaaggagcgaagggtggatgggttaggttatatagggatagacagagatatata
gcaaaagagatacttttagcaagtgttgggaagcggatttcgcaatatttttagctgtgttacagtcgggtgtcgttttgggttttgaagtgcgtcttcagagc
gctttgtgtttcaaaaagcgctctggaagttcttatactttctagaaataggaaacttcggaataggaaactcacaagcgctttccgaaaacgagcgcttccgaaaat
gcaacgcgagctgcgcacatacagctcactgttcacgttcgcacctatatactgcgtgttgcctgtatatataatacattagagaagaagcgcatgtgctgtttat
gtcaaatgtcgtacttattctgccttattgttaggataaagggtgtactgtactgtatcttctgtatatacttccactccatcgggtgtatcgtatgtcttcttcagca
ctacccttagctgttctatgtcgtgcacttctcaattgtgattgtctatcttcaatgtctatattcttctgtattgtgattcataaagaacattattactga
cattaacctataaaaaataggcgatcacgaggcccttctgtctcgcgcttccgggtgatgacgggtgaaaacctctgacacatgcagctccgggagacgggtca
cagcttctgttgaagcggatgcccgggagcagacaagcccgtagggcgctcagcgggtgttggcgggtgtcggggctggcttaactatgcggcatcag
agcagattgtactgagagtgaccatagataaacgacattactatataataataggaaacattaatagacagcatcgtaatatgtgtacttttgcagttatg
acggcagatgtgcagtagtggaagatacttttataaaaaatgagctgttcacattcagtaacattgtatcggagcgttttcttttccggttaagaataaatctg
tgcaaaaaagaagaaggagagggccgaaggaggcggttggcgttggactgtatgaacgtagtagtactgttaagaacacacaaagcagctgtggagtagt
tctgttataatttcacaggtatgtcttggctcattgtgggaagtgtcgggttcgagcacagagcccgcaaatgttctctagttcgtcgtcgtcgtcgtcgt
ggattatagtgtgccaaatagaagaagaacattgacccgggtattgtcaaggaaaattcaagcttgttgaaggcatataaaaaatgttcaggcactccgaa
atactgtgttggcgtgttgcataacactaaggaggatgttttggcttgtgtcaattgattacggcattgatactgtccaaactgcattggaatgagtctgtggca
agaattcaacagagttctcgtggttgcagttatataaagactcgtatttccaaaagactgcaacatactactcagtgcaagcttcacagaacacctcattcgttatt
ccctgttattgcagaagcaggttgggagctgtgaacttttgatgttgaacttgcatttgcaggttgggaagcgagaagagccccgaaagcttaccattat
gttagctgtgtgactgagccgcaaaaaatgttgggtgtgcgtcttagattaaatgtcgcttattgtgtgtatgaagcggaggtgtgtgcaacaaatgtgtgtaaaa
gactctacaaaaatgcaaaattcgtcaaaaatgctaagaataaggtttattactgtagtagtatttttaagttattgttgtgcacttgcgcatcgtcgtgtgaa
ataccgcacagatgctgaataggagaaaaatccgcatcaggaaattgtaaacgttaattttgttaaaattcgcgttaaaattttgttaaatcagctcatttttaacc
aataggccgaaatcgcaaaaactccctataaatcaaaagaatagaccgagatagggttgagtggttccagtttggaaacagagttccactataaagaacgt
ggactccaacgactcaaaaggcgaaaaacccgtctatcagggcgagtgggccactcgtgaacatcaccttaacgaagtgttttggggctgagggtccgtaag
cactaaactgcgaacctcaaaaggagcccccgtatttagagctgttagcgggaaagccggcaacgtggcgagaaaggaagggaagaaagcgaaaggag
cggcgctaggggcgctgtgcaggtgtgacgctgcagctgcgttaaccacacaccccgccgcttaatgtgcgcctgaaagggcgagctgcgcgtccatcgc
ccattcagggtgcgcgaactgttgggaaggcgcatcggtgcggggccttctgcctattacgccagctgtgcgaagggggagtgctgtcgaagcgattaaagt

gggtaacgccagggtttccagtcacgacgttgtaaaacgacggccagtcgtccaagcttcgagctcgagatcccagagctttgcaaattaaagccttc
gagcgtcccaaaccttctcaagcaagggtttcagataatgttacatgcgtacacgcgtctgtacagaaaaaaagaaaaattgaaataaataacgttctt
aatactaactataataaaaaataatagggacctagacttcaggtgtctaacctctcctttcgggttagagcgggatgtgggggaggggcgtaagttaa
gcgtgacataactaattacatgatacctttgtttcgggtgtacaatatggacttctctttctggcaaccaaacccatacatcgggattcctataatcctt
cgttggtctccctaactgtagggtggcggaggaggatatacaatagaacagataccagacaagacataatgggctaacaagactacaccaattacactg
cctcatgtggtgtgtacataacgaactaatactgtagccctagactgtagccatcatcatatcgaaagttcactacccttttccattgccatctattgaagta
ataataggcgcgatgcaacttcttttttttttcttctctccccgttggtgtctcaccataccgcaatgacaaaaaatgatggaagacactaaaggaaa
aaattaacgacaaagacagcaccaacagatgtcgtgtccagagctgtgagggggtatctcgaacacacgaaacttttcttcttcattcacgcacacta
ctctctaagtagcaacggtatagcccttccctccagttactgaattgaaataaaaaagttgccgctttgctatcaagtataaaatagacctgcaattattaatc
tttgttctcgtcattgtctcgttcccttctcctgtttcttttctgcacaattttcaagctataccaagcacaatcaactccaagcttgaagcaagccctct
gaaagatgaagctactgtcttctatcgaacaagcatgcatatttgcgacttaaaaagctcaagtgtctcaagaaaaaccgaagtgcgccaagtgtctga
agaacaactgggagtgctgctactctccaaaacaaaaggctccgctgactagggcacatctgacagaagtggaaatcaaggctagaaagactggaaca
gctatttctactgattttctcgcagaagacctgacatgatttgaaaatggattctttacaggatataaaagcattgttaacaggattattgtacaagataatgtg
aataaagatgccgtcacagatagattggcttcagtgaggactgatatgccttaacatgagacagcatagaataagtgcgacatcatcatcggaagagagt
agtaacaaagggtcaagacagttgactgtatcgcgggaattcttaatacgactcactataggccatattgccatggaggccccggg

SEQUENCE ID NO. 9

gacgcctgcagacagatcccttatagtgagctgtattactgcagagatctaagactgtagatactgaaaaaccccgcaagttcaactcaactgtgcatgtgca
ccatctcaatttcttcatttatacatcggttgccttctttatgttaactatacttcttaagttaacatctggccatgtaacctctgactatagaatttttaaaagacta
gaattaatgcccatcttttttggacctaattcttcatgaaaatatattacgagggtctattcagaagctttggacttctccgagaggtttggcaagtcactcca
tcaaggttctgctgcttctactcttgcgagaaatttacgaaaagatgaaaaagggtcaaatctgtgtgtagatcgtgttgacacttcaataagcgaaattctt
atgattatgattttataataaagttataaaaaaaataagtgatatacaaatttaaagtgactcttaggttttaaacgaaaattctgtcttgtagtaactctttcc
ttaggttcaggttgccttctcaggtatagcagtgaggctgccttattagaccacacctctacggcgatgccgaaattccccctaccctatgaacataitccatttgc
aatttctgctgttcttattatgaattacttaaaagtgttagtacaatactaaaaaagagaatcttttaacgcaaggtatttcttaacttctcggcgacagca
tcaccgacttgcgtgtgactgttgggaaccacctaataaccaggttctgatacctgcctcacaacaccttttaactgtacttcaattggccttaccttctcaggca
gttcaatgacaattcaacatcattgcagcagacaagatagtgggcgalaggggtcaacctatttcttggcaaatctggagcagaacccgtggcatgtgtctata
aaccaaatgcgggtgttctgtctggcaagaggccaaggacgcagatggcaacaaccccaaggaaacctgggataacggaggcttcatcgagatgatac
accaaacatgtgtctgtgtattataatacattaggtgggttgggttcttaactaggatcatggcggcgagaatcaatcaattgatgttgaaacctcaatgtagga
aattcgttctgatgtgttctccacagtttttctcaataacttgaagaggccaaaacatagctttatccaaggaccaaataggcaatgtggcgtcatgtgtgtag
ggccatgaaaaggcgccacttctgtgattcttgcacttctggaacgggtgtattgttactatcccaagcgacaccatcacctatcttcttcttaccaaaagt
aaactctcccaactaattctgcagacaacgaagtcagtaacctttagcaaatgtggctgttagtggagataagcttaaaagagagtcggatgcaaaagtacat
ggcttaagtgtgcgtacaattgaagttcttaccgaggttttagtgaaccttcttgcaggttcaactctgttaccctattagaccaccaacccagccatacaca
aacggcatcaaccttcttggagggttccagcgctcatctggaagggtggacacctgtgcatcgtatagcagcaccaccaatfaatgtatttctgaac
ttgacattggaacgaacalcagaaatagctttaagaaccttaattggcttcggctgtgatttctgaccaacgtgtgtcaccttggcaaacgacgacttcttagg
gcagacattagaatgggtatcttgaatatatatatatattgtcgtaaatgtaaaagtgaaagtgaaagtgaaagtgaaagtgaaagtgaaagtgaaagtgaa
aaacaatagggtccttaataatattgtcaacttcaagtattgtgtatgcaagcatttagtgaacgcttcttattctatgtaaaagccgggttccggcctctcac
cttcttttctcccaatttttcagtgaaaaaggatatgctgcagcgacacctgaattacaacaaaaatttcagtcactgaattgatctgtgcatagcgc
ccctgtgtgttctgttattgttaggaaaaaaaataagggttctgaagagatggaacttgcacttctacatcttgcatacttgcattccacagttgggagatctgcact
ctctagtaggagtaactatgtatcatgtctatagcttcttctgtgtgaattgttattccgtcacaacttccacaacactcagagccgggaagcataaagggttaa
agccttgggtgctaatgagtgaggttaactcaataatttgcgttgcgtcacttccgcttccagctcggcaaacctgttgcgcgttgcgcgttgcgcgttgcgc
ggccaacgcgcggggagaggcggttgcgtattggcgcttccgcttccgctcactgactgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc
agctcactcaaaaggcggttaatacgggtatccacagaatacggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaacgta
aaaaggccgcgtgttgcgcgttttctcattaggtctccgccccctgacgagcatcacaacaaatcgacgctcaagtcagagggtggcgaaacccgacaggact
ataagataaccaggcggttcccccgtgaagctccctctgtgcgtctctgttccgaccttccgcttaccggatactgttccgcttctccttccgtgggaagcg
tggcgcttctcattgacgtcagctgtaggtatctcagcttgcgtgtgaggtgcgttgcctcaagctgggctgtgtgacgaaacccccgttccagccgacgcgtg
cgcttcttctgataactatcgtttaggttgcacacgttgagtgaaagcacacacttaccgcttgcagcagccactgtgtgaacaggttagtcagagcaggtgtatg
agcggtgttcagagagttcttgaagtggtggcctaacttgcgtctacatagaagcagcagatttggatctgtcgtcttgaagccaggttgcacgttgcga
agaggtgtgtagcttctgacccggaacaaacaccgcgtgtgtagcgtgtgtttttgttgaagcagcagattacgcgcgaaaaaaaggatctcaagaa
gatcctttagtcttctacgggtgtgtacgctcagtggaacgaaaactcacgttaagggaatttggctatgagattcaaaaaggatcttcacctatgcctttt
aaattaaaaatgaagtttaataatcaatcaaaagtatatatagtaaaactgtgtcagagttaaccaatgtctaactcagtgaggcacctatctcagcgtatctat
tcgttaccatagttgcctgactcccgctgtgtagataactacgatacgggagggttaccatctggccccagctgtcgtcaatgataccgagacccacg
ctacccggctccagattatcagcaataaacccagccagccgggaaggcgagcgcagaagtggttctgcaacttatccgcttccatccagcttataattgt
ctccgggaagctagagtaagtagtgcgcacgttaattgtttgcgaacgttgttgcctattgctacaggcatcgtgggtgcacgtcgtctgttggatgtgcttca
ttcagctcgggttcccaacgactaaaggcgagttacatgactccccctgtgtgcaaaaaagcgggttgcgttcttctgcgtctcgcgtctgttgcagaaagtaagt
ggccgcagtggttactactatgtgtatggcagcactgtcataattcttacttgcattgcacatccgttaagtcttcttctgtactgtgtgactactcaaaccaagtc
tctgagaatagtgtatgcggcgaccgagttgtcttgccttgccttgcataatccgggataataccgcgccacatagcagaactttaaagtgtctcatattggaaa
acgttcttccgggcgaaaacttcaaggtatctaccgctgtgtgagatccagttcgtatgaacccactctgtgccaccaactgatcttcagcatcttttacttacc
agcgttcttgggtgagcaaaaacagggaaggcaaaatccgcgcaaaaagggaataaggcgacacggaaatgttgaaatcactaactcttcttcttcaatatt
attgaagcattatcagggttattgtctatgagcggatacatatttgaattatttgaaaaaataacaaatagggggttccgcgcacatttccccgaaaagtgc
acctgacgtctaagaaacattatcatgacataacctataaaaaataggcgtacacgagggcccttctgctcgcggttccgtgtgagcgggtgaaaaact
ctgacactcagctcagctcccgagcgttcacagcttctgttaagcgtggtatccggagcagacagcccgctcagggcgctgcagcgggttggcgggt
ctggggctgtgcttaactatgcctgactcagacagcgttctgtatgaggtcaccataacgcttaagcataaacacgactatcggcttcttctcgtatgtat
atatatacaggccaacacgcagatataaggtgcagctgaaacagtgaagctgtatgtgcgagctcgttgcatttctgggaagcgctcgttcttgcgaacgctt
gaagttcttattccgaagttcttcttctagctagaagatagtaggaactcagagcgcttggtaaaaccaaaggcgctcgtgaagacgcacttcaaaaaacca
aaaacgcaccggactgtaacgaagtactaaaatattgcaataaccgcttccacaacattgtctaaaagtatcttctgtatatactctgtctatattccctata
taacctaccatccacttctgctccttgaacttgcacttaaaactgcacttcaatttttattgttattcttagtatacaaaaaaattgtagtaagaaga
tattcatagagtgaatcgaaaaacatacgaaaatgaaacatttctatactgtagtataagacaaaaatagaagaacccgttcaattttctgaccaatgaa
gaatcatcaacgtactaccttcttcaacaaagtgcgaactccacatcgggtatagataataatcgggagtgcccttattctgaaaaaatgaccccgagctt
cgctagataatcagctaacccgggaagtgtggaactcagggctttttttatggagaagaaaaatagacacaaatagctccttcaacttcaacgagctacagctg
aaaaagtatacaagagactgcattatagcgcacaaaggagaaaaaaataacttaagtcttcttggtaaaaaatagcgtctggatcagcttattttgtaga
acaaaaaagaagtatagattcttgggttaaatagcgtctcgttgcatttctgttcttgaataatgcagctcagattcttggtaaaaaaatagcgtctgc

cgttgcatTTTTtttcaaaaaatgaagcacagattcttcgttgtaaaatagcgcttgcggttgcatctctgtctgtaaaaatgcagctcagattcttgttgaa
aaattagcgctctcgcttgcatTTTTtttctacaaaatgaagcacagatgcttcgttgcttgcatgcaactctttctttttttctctctccccggttggtgtct
caccatatccgcaatgacaaaaaaatgatggaagacactaaaggaaaaaataacgacaaagacagcaccaacagatgtcgttgccagagctgatga
ggggtacttcgaacacacgaaacttttcttcttcattcacgcacactactcttaatgagcaacggtatagcgcttcttccagttacttgaatttgaataa
aaaaagttgcccgttggctatcaagtataaatagacctgcaattattaatctttgttcttcgtcatgttctgttcccttcttcttcttctttttctgcacaatatt
caagctataccaagcatalcaatcaactccaagctttgcaaatgataaagcggaattaattcccgagcctccaaaaagaagagaaggcgaattggg
taccgccccaattttaatcaaatggaatattgctgtagctcattgtccttcacttcactaacagtagcaacggctccgaacctcatacaactcaaacaaa
ttctcaagcgcttccacaaccaattgcctccttaacgttcatgataactcatgaataatgaatcacggctagtaaaattgatggttaataatcaaacccac
tgtcacctgggtggacggacaaaactcggtataacggttggaaatcactacaggatgtttaataccactacaatggatgatgtataactatctattcgatga
tgaagatacccccaaacccaaaaaagagatcgaattcttaatcgaactactataggcccatggacgaagaatccagttcattctatgtacctatgctg
agaatcgtgccaataagaagccaatacttcttagatgatgcaataatattaaaaataaaacaaaacagaaggctg

pGN205:

SEQUENCE ID NO. 10

ccgggtggaccgggccccctcagggtcgacggatcgataagctttcgtcattgaaaagaaggataagaatggacgatgggaagaagctctcgtt
gttcaggagatcagaaaacagcaactgttccaaatctaaggaggagagaagaataatcagaatttcgtctgtaacaaggctggaaactggaga
tccttcgatccttcgatcgtgtgttgccgaagccaagaacccctgtccaaagaattcatcgtgaagatctttctgatacaactgtcaaggctggagccac
tctcaagttcattgttcattatgttggtgagccagcaccagatgtaacatggcattcaatgaaaaggaaatcgagagagcaaggctcaaatgaaaa
tgagccatacatcgcgagatttgccttccaaaggcacttcgtaagcaaatgggacaaatcgaggtaactgatgtcttcgaagatcgtgcaactcttgactggaaac
actatcaatatcaaggtaaaaaagcaagccaacgaacccaagggaacaaatcgaggtaactgatgtcttcgaagatcgtgcaactcttgactggaaac
caccagaggatgacggaggagagccaattgatttatgaaattgaaaagatgaacaccaaggacggaatctgggtccatgtggactgtagtgagg
atacccacttcacagtcgattcacaagaaggagatcattacaagttccgtgtcaaggctgtcaacagcgaaggaccttctgatccattggaaactg
aaaccgataatttggctaaaaatccattgatcgtccagatagaccaggctcgtccagagccaactgatgggattctgatcatgttatcgaatgggat
ccactagtctgaagcgctgtaagggggcccctcgtcgtgagtcgtcacaatcacctgaaactccaaaggcagccagtgaggaaactggaagaaga
agaaaaagatcatcgaacagggttattttcttctgtcaaaaaagatgaattattgatttcagccagatactccaaaactagcagcgagaaggtct
gcaagtcgttcacagtcgccagagaatcgccggaagtgaagcaagaggatgtttttcaaaaatcaataactgatcataattttatgttggtaattt
aagaaaaataatgaaatccctcgtgaattatcaagattgcagatattaatttcgagaaaaatgagataatcagagctattgttaaattttctgatttcag
actgaaacttcggaatcaagagaaaaatcaaaagaaaaggatgacgggagatgacgctggcacaccgaacagctatagaagccgggaaacttc
accagctccaaaagggtccaaggagaccagggttgcataaagcttctcgtgattaatctcatttcaattttcagagaatcagagctcctgaaaaatcc
ccggttcgtcaagatctccagaaggcttccagcacgttccccgtcacgatctctagacggcgccgagaagaagctcagaaagaagaacaaatccg
aagagccagcaccgctaccagagaaaaagaaagagccgctggatattctacgaacaaagacccggaggagcatatattccaccgccaaactt
cgacttatgcaacacagattagtgataagcaaatgaacagatcagagaatgaattgggaaagaatgaagaaaaagattcacgattggtaacag
agtcaacgcgaagaatcttgcataattgtcagagaacttctcaagagaatgtgattcgtcaaaagtgaagtgaagaaatcgaaggaaggaagaat
taatttaattttcaggggacttctcgtcgtgacattatcaagctcagggttctcaccaggattctcactacgtctatgcagcttggggcgattatcaac
tcgaaatccctcatgctcgtgaaacttctcgtcgtcgtgattgtacagttcaaaagaagtttcgtgaaatgacagagggcgctacgggtgaactgtat
caaatctacgcacatttgattatcaacaagttgctcacgaagtctcgtcgtgaaatcatgattctgagcttggaagcaactgatgattcagttga
agtcgccaattcgttctgaaagagtggtgagcaaaagcttggagattgctccagcagctcttaacagtgctacgaccgtcttcgtcaattctcag
gaaactgaaagatcggaataatgcactggatcgacgtattcagtatatgattgagactgcaatgcagattcgaaggacaaatttgcgtgaaggtagaa
tatataaatagtttatgaaaaataaattagaataatttaaaatcttactatgccaatcaggcgaccttttgcgcatagtctattatgaaaaatttggag
aatttctatattctcgtcgtgaaatcggaaatcgacgagatcttctgcttctgctcagctcgtcttctgctcccttctcgttcttctgtgtaca
ccaagaaccttgtgattcatcaactgaatctgtgactggttgtgtcactggatgcactagacgactgattctcgagaatcagattgattgctgatt
agggtagcttagaaattgggaataatcgaactttgaaaaatattcaggaggatgaaaaaatattctcgacaatctcacaatttactattgcaccatgt
tgctccaacattttcattaaaaatgaataaagtgtagaaatcggaatttgcgaattttcagaccatttttaagcattttcaaaaaaaatgcagctga
ataaattgctattttcagataaatcgagcgattttctgttctgacactagtttttagtttaaaaaatgttgaagaacattggtgcaatggttaattcag
aatttcatgtgttttttcaataaacaattatccaatcttccaaactcacatttgcggagctgggctatcaagaatctgctgctgttttaagacgagc
atctctgatactgaaaaatatttttaataaaaactgaatacaactaaaccacttatttaacttctcgaatcttctgcttctggttcgtacgatgacggtga
gaagccaattgtagtagttgattgttcaagcttctggtgtgtacgtcagtgctcgtcaatgctatttagttataacttaggcctaaagattcaattaatg
aagtattaaatttcttgaacctttaaagatgatttttgattagaacatataagacaggtttacctatctataaaaaacagatcaaaatagatagc
accaaatcggaataatccatgcttaccctggcatctaggaacgtgttctagaagatttctacgtaatcgtatgaagaataaacaatttgcgttggccag
caaaaataggggttaagtggtgagatgtttttatagctaacccgaaaattttatgttttttgcagaagaaccactgaaaaacccctaatgtatatacttt
ttggagcagcttctgcttctttttagcaataaaatcgataaaacagaatttaagtgtaattgttcacatttagtttctattttatcaaaatttgcctcaaaaa
cttcgaagcttctcctcaaaaaatgcttaaaaaagggttttcagttgttttcaacttaaaaaagcttaatttaactaaaaatccatattttccaactttg
tcacaacaataaaatgctgttcaaaatgtgttcgaaaaatgttttttttaatttttaaaaaatagttttcttctgctgggacacatacatttttgggc
gtaattttcagttcaatttccattttacaaacataatcataagctacgtctgatctctcgcacttacctgcgctgattcgaagaacaaccgtagc
caaaagaacaagaagaacagcagtagttgtgtagtggtgacgttcacgcgaatactgaccaatggctgtgggtctcacttccgtactattgaga
gaggggagactgaagatggcaattgaggacagtgcttctgacgcagcagcatccataagcataatccaggagggtgagagaaaaacttctgtt
ctaagccctctccttgaatacacaatataataccgaagaatggcctaattgaatggagctcagctgttctgtgatttccaaggcatcatcagtg
aaataactgaaagaagaataaataatttgcaggcgatccggcggtcattgaagacttggacttgaggaggaggatcagatcatccataca
cttaatttggaggatgcggtgatccgaaaatggccttagtaagtactgaccacacgcgggggcttaatttaaaatgaattccatttcagatgt
gttcaacttagatccagaattcgaagaacagcagggtttatgaggagatccgtaaggaaatcattggaacgcggaatttccgatgaggatgggt
gcgacgagttggtatgaagaagagggtagtgatgtggaagaggctccgaagaagactacagagattattgataatcagatcagaattgactgctt
tcagaaggatattcattttagtttggccggcaaatctgaattgtccggttgcggaatttgcgtgaatttgcggaaaaaaaatccggaatttattta
aaaactttttgtaaaaaatataaattgaatttgcacattttagagaagctacgtgacaatgcatactttagactaccaagaagctgtctacaaattgctg
aaaatgaagattccagacagcatgcaggtcagcgatgttgcagaagaatttgcacaaaaaaacaaacaaatcataaaatttaaaaaaaactcc
gttttttctttttttataggaaaaaacaaaaaatgtatttttccaaatttcaaaatctatccccgaaattttcaattttcttcttcagaacgaactctg
cgcgagctgtcgtgattgtgtcacaacgcgtacctacgagcgattctacggaatgctcatcgaaacgttttgcgacttgcctcgaataccagcaa
tactttgaaaagctctgcaggacagctatttccagattcaccgaattgacatcacaacactgcggaatttggctgccttattgtctcattgtctcagc
gatgctattgactggaagatttggccgatatgaaaatgaccgaaggagacaaacttcttgcgagaatctataattaaatataatttaattgaacttggc
gaggcgatgggaatggttaacttcatcagagattactgactcgtgagtttccatagagagattgttttctgatttcaattttcttcagaaacttgcct
cattgttttggatttccacgaactaaatccgaacgcgcagcagatttgcagatcaacttctcacaatgattgattgggtgtgtgacgttgaacttgc
tgaaatggctggcaagggtctcaagaagaagaagggaatgtggtgatcagttgaaggccgaatcaagctcagattcatcgtctcgttctcggattcgtca
gactcgtctgatttctcgatttgcagatttccgactcgtctcagatttctcacttctcagaatcagagccagaaccaccgaagaaaaagaagaa

gaagaacagtgagagaggtccaaaaagaaggaaaaagagaatttggtcgacgggacgtggagacaagagagctgaacgtcatcgatcaaa
gtgtggagaacaaggacaaggatcgacgacgagcaggaattcgacgaaaaatcgccgagacgagagatgaccgcaaggatcgagtaa
agatcgctcgctcaagactcggatgatgagatcggaagggtcggaacgctgggaagattcaggggaaagacgtcgcgagatcggtatcgac
gtgatcgaacaaggatcaggagggatcacctgaagatcgccgtgacggaagcaaggatcgtaggagatcgacgtgatcgccgtgatgactctg
atgatcgatgtaaaactcgctgggatagaagtgaagagcaggaggacgtcgctggaagtggaaatcgatcgacgcccagctcggtgaattt
caaattttaataactgaatttggtttttctattattatttctcttggttttttcttcttcttaaaaaaattcaatccaaatcaaatgagcggtt
tttttcttcttccgtctccaattcgattccgctcctctcatctgaacacaaatgtgcaagttaatttctctcgtttcatttcattaggacgtgggggaatt
gggtggaagggggaacacacacaaaggatgatggaatgaaataaggacacacacaaatgcaacaacattcaattcagaataatggaggaggttaa
aagaaacataaaaaatatagaggaggaaggaaaactagtaaaaaaagaagaattaggcgaacgatgagaattgctcgttggcaaatg
cgaatccgtatggagagggcacgttggcgaaggcaaatgttcggtatgagatctgtaaaaaattttaagtgaatttgggtgtcttcttaaaaaattt
ccgatttctcgttgaaattacgggtccaggctcgcacacgttccaattttcaaatcaaaagagccttaattggcgtgtagtctgaatttctcgttttga
aaatttttctccgttaaatgaaattgatgtattttatgatgtttcaataaattcaagaactgggtgaaaactcggaaaattgtaactacagtaatcca
atccttaaggcgacacacttttaattgctccccaatcacgataatttttaagattcgtagagcgcccgccaccggtgagctccaattcgccct
atagtgcgctgattacaattcactggccgtctgttaaacgctgtagctgggaaaacctggcgtaaccaacttaacgcttgcagcacatcccc
cttcgacgctggcgtaataatcggaaggcccgacccgacccgaccccttccaacagttcgtagcctgaattggcgaatggagcgccctgtagcg
gctgattaagcgcggggtgtgggtgttacgagcgagctgacccgtacacttgcagcgccctagcgcccgctcttctcgttttcttcttcttct
cgccacgttgcgcttctcccgtaagctctaaatcggggctcctttaggttccgatttagtgcgtttagcgcaccccaaaaaacttgat
taggggtgaggttcacgtatgggcccacgcccgtatagacggttttgcctttagcgttggagtcacgctttaaattagtgactctgttccaaact
ggaaacaaactcaacccatctcggctattctttgattataagggaatttgcggaatttgcctattggttaaaaaatgagctgatttaacaaaaattaa
cggaatttttaacaaaataaactgtaacgtttcaatttcaggtggcacttttggggaattgtcgcggaacccctatttgttttttctaaatacattcaaatat
gtatccgctcatgagacaataacccgtataatgtctcaataatgtgaaaaggagagatgagtagtattcaacatttccgtctgcccctatttcttcttctt
cggaatttgccttctctgtttgtctaccagaaacgtgtgtaagtaaaagatgctgaagatcagttgggtgcacgagtggtttacatcgaactgga
tctcaacagcggtgaagatccttgagagtttgcggcggaagacgttttcaatgatgagcacatttaagttctgtagtgccggttatatccgatt
gacccggggcaagagcaactcggctcgccgatacactatttctagaatgacttggtagtactaccagtcacagaaaagcatcttaacgagtggaat
gacagtaagagaattatgcagtgctgcataagcatgagtgataacactcggcccaacttacttgcacacgagtcgaggacggaggaaggaaccc
gcttttttcacaacatgggggatcatgtaactcgccttgatcgttgggaacggagctgaatgaagccataccaacgacgagcgtgacacacagat
gcccgtagcaatggcaacacgttgcgcaactattaactggcgaaacttacttacttagcttccggcaacaataatagactggatggaggcggaata
aagttgcaggaccacttctgcgctcgcccttccgggtggtgttattgtctgataaacttgagccggtgagcgtgggtctcgcggtatcattgcag
cactggggccagatgtaagcccctccgtatctgtatctacacgacgggcagtcaggcaactatggatgaacgaaatagacagatcgtgagat
agggtccactgattaaagcattgtaactgtacagcaagtttactatatacttagatgtatttaaaacttcttttaatttaaaaggatcctaggtgaa
gatcctttttgataatctcatgacaaaaatcccttaacgtgagtttcttccactgagcgtcagaccccgtagaaaagatcaaaaggatcttctgagatcc
tttttctgcgctgaatctgctgttgcacaaaaaaaccaccgctaccagcggtgtgttgcggatcaagagctaccaactcttttccgaaggt
aactggtctcagcagagcgagataccaaatactgtcttctagtgtagcgttagtgccaccacttcaagaactctgtagcaccgctacatactt
cgctctgtaactctgttaccagtggctgtgtccagtgccgataagtggttcttaccgggttggactcaagacgatagtaccggataaggcgacg
ggctcgggtgacgggggggtgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagcattgagaaaag
cgccacgcttcccgaaggagaaaggcgacaggtatccggtaagcgccagggctggaacaggagagcgcacgaggagcttccaggggggg
aacgctgtgtatctttatgctctggttccgacctctgactgagcgtcgtatgttctgtagctcgtcagggggggcgagcctatggaaaaacg
ccagcaacgcgcccttttacggttcttggccttctgctgcttctgctacatgttcttctgctgtatccctgattctgttgataaccgtattaccgct
ttgagtgagctgataccgctcgcgacccgaacgaccgagcgcagcagtgatgagcaggaagcggaagagcgcccaatcacgaacccg
cttccccgctgttgccgattcattatgcagctggcacgacaggttccgactggaagcgggcagtgagcgaacgcaattaatgtgagta
cctcactcattaggcacccagccttaccattatgcttccggctctatgttgtgtggaattgtgagcggatacaaatcacaaggaacagctatg
accatgattacgcaagctcgaattaaccctcactaaagggaacaaagctgggggggagcttcccaaatcgttctccgcttgaaaaacgaaagt
ggaccttgacatccgaaaaatggcgaaaaatgaattgagcttttgggtcgaaaaaatgttttagaatgctgagaacacgttaaacacgaag
atcatattttttgagaccggatgctctgaaaaatgtctgacatagatttaaaaaagcatatataatttttcaattcaacgtgaaagtgttgcactttata
gaatctctatttggacattgttttttttaactgaggcagtttttgaacaccttttgaacattgaatcttgaagtatactgtcgaaggactgactga
cggttcgaatgccagaagaaaaataatttgaatctcgcgtaaatgagaaatcaaccgctccactggacaattggaaaaaaatttattcggga
ggcgacaacggtatttgcgaattgtgtatttctcatttttataaattcttcttattatctgctgttgtgagaaatttaattgtattcaacttttt
atagtaagata

SEQUENCE ID NO. 11

ccgttggtaccgctagccgtacgaacccgggttctagaactagtggatccaccattgagatcaacatgatcagaatcccaatcagttggctctggacga
ccgtggctatctggacgatcaaatggatttttagccaaaatcgtgttgcattccaatggaatcagaaggctcttgcgtgtgacagccttgacgagaa
ctgtaatgatctccctgtgtgagtgaatcgactgtgaagtggtgtaicccactacgtccacatggaaccagattccgtcttgggtgtcatctttcaatt
catagaactcaattggctctccgcctcctgtggttgcctcagtcgaagatgacagttacctcgattggctcccttgggtt
cgttggcgttcttttacccttgatattgtagtgaactgtcagttccattaatgttgggtgcagtgatggtatattttccactttgtctacgaagtcctttggca
aagcaaatctcgagatgatgtggtcattttcaattttagccctgtctctccgattccttttccattgaaatgaccatgttacatctgtgtgctggctccacatca
atatgaacaatgaacttgagagtggtccgaccttgacagttgtatcagaagaatcttcacgaigaattcttggagcaaggtttcttggcttcgaacaac
acgatcagaagaagatcagaaggactcgcagctccagccttgttacgagcagaatacttgaattgatacttctccctccttaagatttggaaacagttgtgt
tttctgattctcctggaacaacgagagctgttcccatctccactgtccattcttccattcaatgacgaagaaagcttatcagctcgaaggtttgtatttct
cctcgtcgatgtggctcacaatcaccttgaacctccaaggagtcacgttcaggaaacgtgaagaagaagaagaagagagctatcgaagcaggtttgtatttct
tttggctcaaaaagatgaaattattgatttcagccagatactccaaaactagcagcgagaagcttgcgaattgttcacagtcccgacagaggaatgtcgcg
ggaagtgaagcaagaggatgtttttcaaaaatacaataactgatcataattttatgtttgttgtaatttaagaaaaataatctgaaaattccttgaaattatc
aagattgcagttatatttcgagaaaaattgagatatcatagagctattgttaaattttcttgatttcagactgaaacttcggaaaaatacagagaaaaatacaa
aaaaaggatgacgggggatgatcagcttggcacaccgaacgctatagaagccgggaaacttcaccagctccaaaaaggtccaaggagaccaggt
ttgtcaaaagcttcctcgatatttctatttcaatttttcagagaatcagatgctctgaaaaatcccggttcgtcaagatctccagaaggtcttcag
cacgttcccgctacgatctccttagacggcgccgagaaagaagtcagaaagaagaacatccgaagagccagcacccgctaccagagaaaaagaa
gaaagagccgcttgataattctacgaacaagaacccggaggagcatatttccaccgccaaacttcgacttatgcaacaacagattagtataagcaa
agtgaacagttatcagagaagaattgggaagaagatgaagaaaaagttacacgattgttgtaacagctcaacgcgaagaacttgttcaaatgttcag
agaactttccaagagaagtgtgattgttcaaaagttagtgtaagaaatcgaagaaagaagaataatttaattttcaggggactcttcctcgctgac
attattcaagctcaggtcttctcaccaggtattctctacgtctatgcagcttggcgccgagttatcaactcgaataatccctatgttgcgtgaacttctcc
gtcgtctgattgtacagttcaaaagaagtttccgtagaatgacagaggcgctacggtgaacgtgatcaaaattcatcgacatttgattaatcaacaagtt
gctcacgaagttcttgcgctggaatatgatattctgatgcttgaagaaccaactgatattcagttgaagtcgccattgcgttctgaaagagtgtggag
caaaagcttctggagattgtctccagcagctcttaacagtgctacgaccgtcttcgtcaattctcatggaactgaaagatcgaaaaatgcactggatgc
acgtattcagatatattgtagacgtcgaatgcagattcgaaggacaatttgcggtgaagtgagaatatataaattagttattatgaaaaaaataaattaga
ataatttaaatcttactagccaatcagggcgaccttttgcgcatagttctatttgaaaaatttggagaatttctcatattctgctcgaaatctggaattc
gacgagatcttctggcttctgtgcagctgcacgtcttgccttcttcgcttgcgttcttctgtgtacaccaagaaccttgtgagttcatcaactgaactctgt
gactgctgtgtgtctacgtgtgactgactgacgtatgcttgcacgtacgaatcagattgattgctcgattagggtgcactgaaaatttgggaataatcgaact
ttgaaaaatttcaggagaggttaaataaaaatttctcgacaactctgaaaatttacttgcacatttcttcccaacttttcaataaagttaattgaaaaa
atgtagaaaaatcggaatttggcaatttcagaccatttttagacatttcaaaaaaaattgcagctgaaataaatgtcttttcagataaatcagcgcgattt
ctgtgtctgacactagtttttagtttaaaaaatgttggagaacatlggtgcaataaggtaatttcatagaatttccatgtgttttttcaattaacaaattc
aaatcttccaaactcacatttgcggagctgggctatcaagaatctgtctgcagtttataagacgagcatctctgatctactgaaaaatatttttaacaa
aactgaaatatacactaaacccacttattaactttctcgatcttctgtctgttcggtacgatgacggtgaaagccaattgtagtgtgatttggttcaagtc
cttctgggtgtgtacgtcagtgctctgcaatgctatttagttataactttaggcctaagatcaatttaagtgattaaatttcttctgaaaccttctaagat
gatcttttggattagaacaatataagacaggtttacctaatttataaaaaacagatcaaaatagatacgaacaaatcggaataaccatgcctacctggcat
ctaggaacgtgttcttagaagatttctacgtaatcgtatgaagaataacaatttgcgttggccagcaaaaatagggtttaagtgggatagtttttat
tgtaaacccggaataatttatgttttttggcaagaaccactgaaaaacccctaattgtatataacttttggcagcgtctgtgtttttagagcaataaaa
tgtaaaaaacagtaatttaagtgttaattgttccattgtatttcttcaaaatttctgtctaaaaaacttcgaagctgtctcaaaaaatgattaaaaaa
aggggttttcagtggttttccattaaaaagcctaattttaaactaaaaatccatcatatttccaactttgtcacaacaataaaatgctggctcaaaattgttgc
aaaaaatgttttttttaatttttaataaaaaatgttttcttgcgtggacacatacatttggcgctaatttccagttcaaatcttatttacaacat
aatcataaagctacgtctgatctctctgcacttacctgcgctgattcgaaagaacaaccgtagccaaaagaacaagaagaacagcagctagtgtg
ggtagtggagcttcatcagcaatctgaccaatgtgtgtggggtctcatttctgactattgagagaggggagactgaagatggcaattggaggaca
gtgtcttcgacgcacgcatgatccataagcataatccaggagggtatggagagaaaaatcttgttctaagccctccttggtaatacatcacatatct
aataccgaagaatgctaattgaatggacgtcagctgtgtgtgtgttgcgaaggcatcagatgaaataactgaagaagaagaattaaataattattgc
aggcgtaiccggtcattgaagacttgacgttattgaggaggaggatcatcatcacttaatttggaggatgcgggtgatccggaaaatg
ggcttgtaagtgtgactgcacacgcggggggcatttaataaattgaatttccatttcagatgtgttcaactagatccagaattcgaagaagaacga
ggagggttagtaggagatccgtgaaggaaatcgtgaagccggaatcttgcgttgatgagtggtgtggcagacgttgatgtaggaagagggtgagt
gatgtggaagaggctccgaagaagactacagatatttattgataactgatcagaattgactgtcttcagaaaggtattcattttagtttgggtccgcgcaa
atctgtaagtgtccgggttgcgaaaaattgtgaatttgcggaaaaaaaatccggaaatttataaaaaacttttgaataataaatttgaactt
ttcagagaagcttactgacaatgcaatcatttggactaccaagaagctgtctcaaaattgtgaaaaatgaagattccagacagcatgcagggtcagc
gatgtgtcaaaagaaaaatttcgacaaaaaaaccaacaatcataaaatttaaaaaaaactccgttttttcttttttatacagagaaaaacaaaaaa
atgtatttttccaaaattcaaaaatactatcccgaattttcaatttttcttccagaacgaactctgcgcgatgtgtgtcagttgtgtctcaacagcgta
cctacgagcgatttctacgaatgtcatcgaacgttttctccgacttgcctcgaataaccagcaatcttgaaaaagcttgcaggagacagctattccac
gattaccgaattgacatcaaaaaactcgggaatttggctgccttattgtctatttgcctgcagcgatgctattgactggaagatttggccgatgtaa
aatgaccgaagaggacacaacttcttctggcgaatctatatttaataatatttaatagaacttggaggcgatgggaattgttaacttctcagagag
ttactgtacgtgattttctagagaagattgttttcttaatttccatttccattcagaactttgttcgcatattgtttgttatttccacagcaataatccg
aacagcgacgatttccgaacttccaaacttccaattgattgggttgggttgcagttggaacttcgtgaattgttggcgaagggttccaagaagaaga
agggaattgttgatcagttgaaggccgaatcaagctcagattcatctgtcttccgacttcagtcagactgtctgatttctcggatttgcattcagctaccc
gactgtcttcagatttctatcttctcagaatcagagccgaaccaccgaagaaaaagaagaagaagaacagtgaaagagatttccaaaaagaag

gaaaaagagaataattgtctgacgggagatctggagagcaagagagctgaacgtcatctgtatcaaaagtgtggagaacaggaacaggaatgacgtgacg
tcgccaggtatctgacgaaatctgctggccagaacgaggagatgaccgcaaggatcgagtaaaatgctgctgctgaagctggatgatgaggg
atcggaagggctgctgaacgtcggggaagattcaggggaaagacgtcgccggagatcgggatcgacgtgatcgaacaaggatcaggaggatcaccg
tgaagatcgccgtgaccgaagcaagatcgtgaggatcgacgtgatccgctgctgatgactctgatgatgatcgtaaaaactcgtcgggatagaagt
gaagcgcgaggaggacgtcgtcgtgaagtggaaatcggtatcgacgccgacgtcgtgaattttcaaaatttaataactgaaattgttttttcttatt
attattattctctttgttttttctcgttttaaaaaaataatctcaaatctaaacatgagcgggttttttctcttccgtctccaattctgattccgct
ctctcatctgaacaatgtgcgaattttattattctctctctttatttaggacgtggggggaaattgtgggaagggggaaacacacaaaaggatg
alggaaatgaaataaggacacacaatatgcaacaacattcaaatcagaataatggaggaaggttaaaagaaaaataataatagtaggaggaa
ggaaaactagtaaaaataagcaaaagaaattaggcgaaacgatgagaattgtctctgtggcaaatgtgcaatccgtatggagagcgacgttttggcga
aggcaaatgttcggtagggatctgtaaaaaatttaagtgtgaatttgggtgttctcttttcaaaaattttccgattttcggtgaattacgggtccaggct
cgacacgtcttcccaattttcaaatcaaaagagccttaaatgggctgtagtgtcaatttctgttttgaaaatttttctccgttaatacgaatttgatgtatt
tattatgattttcaataaaattcaagaaacgtggtgaaaactcggaaaattgtgaactacagtaatccaatctctaaaggcgacacacctttaatgtccgc
cccaatacgatatttttaagatctgtagagcggccgcccacggcggtggagctccaattcgccctatagttagtctgattacaattcagtggcgtctgt
ttacaacgtctgtgattttggaaaacctctggcgttaccacaacttaacgtccttgagcagcacatcccccttcgccagctggcgtaatagcgaagaggccc
gcaccgctgccttcccaacgttcgtagctgactgaatggcaatggagcgcgccttgcgttcgcccataagcgcggcggtgtgtgtgtacgc
gcagcgtgaccgtacacttgccagcgccctagcgcccctcttctcttcttcttcttcttccagcttcccgctcccggtcgaagctca
aatcgggggctccttttaggtttcgatttagtgtcttaccgcacctcgaccccaaaaactgattagggtgtgtgttccagttgtggccatcgccct
gatagacggttttgcctttgacgttggagtcacgtctttaaattgtagactctgtttcaaaactggaaacaacactcaacctatctcgggtctattctt
gattataaggagatttgcgatttcggcctattgtgtaaaaaatgagctgatttaacaaaaattaacgcgaatttaacaaaaataaacgtttacaatttca
ggtggcacctttcggggaatgtgcgcggaacccctattgtttattttctaatacatctcaaatatgtatctcgctcatgagacaataacctgtataatgc
ttcaataatattgaaaaaggagagtagtagtattcaacattccgtgtcgccttattccctttttgcggcattttgccttctgtttttgtcaccagaaac
ctggtgtgaagtagaaaagatctggaagcatcattgggtgtcagcgtgggtttacatgaacttgaatctcaacagcggtaagatccttgagagtttccg
ccgaagaacgttttccaatgatgacgtctttaaagtcttctgtgtggcgcgttatctccgattatcagcggcgcaagagcaactcggctcgccgat
acactattctcagaatgacttggtagtactaccagtcacagaaaagcatcttacggatgtatgacatgaagaaattatgcagtctgccaataag
catgagtataacactcggcgcaacttactctgacacgacgtggaggaccgaaggagctaacccgctttttacaacatgggggcatgtataactcg
ccttgatcgttgggaaccggagctgaaatgaagccataccaacgacgagcgtgacaccacgatcctgtagcaatggcaacaacgttgcgcaact
attaactggcgaaacttactctgacttcccgccaacaattaatagactggaatggaggcggaataaagtgcaggaccactctcgtcgtcggcccttc
ggctggtcgtgtttatgtctgataaacttgagaccgggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtgaagccctcccgatctg
agtattctacacgacgggcagtcaggcaactatggatgaacgaaatagacagatcgtgagatagggtcctcactgattaagcattgtgtaactgtcag
accaagtcttcatatataacttttagattttaaaacttcaattttaaaaggatctgtgaaagtacttttttgcgtaactctgctgttcgcaaaaaa
cgtagtttctgttccactgagcgtcagaccctgtagaaaagactaaaggactctgtgagactcttttttgcgtaactctgctgttcgcaaaaaa
aaaccaccgctaccagcgggtgtgttttgcgggataaagagctaccaactcttttccgaaggtaactgtcttcagcagcgcgagatacaataact
gtccttctagtgtagcgttagttaggccaccacttcaagaactctgtagcaccgctacatcctcgtctgtataactctgttaccagtggctgctgccag
tggcgataagtctgtcttaccgggttgactcaagacgatagtaccggataaggcgcagcggctcgggctgaacgggggggtctgtcacacagcg
cagcttgagcgcaacgacctacccgaactgagatacctacagcgtgagcattgagaaaagcgccacgcttccggaagggaagaaggcggaacagg
tatccggtaagcgcgagggttggaaacgagagcgcacgaggggagcttccagggggggaacgcttggtatcttatagtcctgtcgggttcgccac
ctctgactgagcgtcgtattttgtatgctcgtcaggggtggcgccgctatggaaaaacgccagcaacggcctttttacggttttcggttcgtggttcgt
ggcctttgtcactatgtcttcttccgtcgtctacccctgaggttctgtgaaacgctattaccgcttttagcgtgactgataccgtcgcgcagccgaacga
ccgagcgcagcgagtcagtgcgaggaaggcggaagagcgccaatacgaacccgcttcccgccgttggccgactcttataatgacactgc
gcacgacagggtttccgactggaagcgggcagtgagcgcacgcaataatgtgagtacccactacattaggcaccccggtttacattatgtct
tccggctctatgttgtgtggaattgtgagcggataacaatttcacacaggaacagctatgacatgattacgcaagctcggaaataacccctactaa
agggaacaaaagctgggggggactctccaaaactcgttccgctctgaaaaacgaaagtggaccttgacatccgaaaaaatggcgcaaaaaatga
aatgagcttttgggtcgaaaaaaatgttttaagaatgctgagaacacgttaaacacgaagatcatattttttgagaccggatgctcgtgaaaaatgtctg
acatgatttaaaaaagcatalatatatttttcaactcagtgaaaagtgttgcaactttaagaatctcctattggcacattgtttttatctaactgaggcag
ttttgaaacacttttgaacatttgaactctttgaagtatactgtcgtgaaaaagactgacttgagcgttcggaatgccagaagaaaactatatttgaatctgc
gctaaattgagaaatgcaaccgcctccactgacaaattgtgaaaaaaatttccggagcgcaaacgtatttccgaaattgattttctgtgtatttct
catttttaaaattcttcttgattatctgtctgtttgtgagaaaatttaattgattacaactttttatagtagata

Claims

1. A method of identifying DNA responsible for conferring a particular phenotype in a cell which method comprises
- 5 a) constructing a cDNA or genomic library of the DNA of said cell in a suitable vector in an orientation relative to a promoter(s) capable of initiating transcription of said cDNA or DNA to double stranded (ds) RNA upon binding of an appropriate
- 10 transcription factor to said promoter(s),
- b) introducing said library into one or more of said cells comprising said transcription factor, and
- c) identifying and isolating a particular
- 15 phenotype of said cell comprising said library and identifying the DNA or cDNA fragment from said library responsible for conferring said phenotype.
2. A method according to claim 1 wherein said
- 20 library is organised into hierarchical pools prior to step b).
3. A method of assigning function to a known DNA sequence which method comprises
- 25 a) identifying a homologue(s) of said DNA sequence in a cell,
- b) isolating the relevant DNA homologue(s) or a fragment thereof from said cell,
- c) cloning said homologue or fragment thereof
- 30 into an appropriate vector in an orientation relative to a suitable promoter(s) capable of initiating transcription of dsRNA from said DNA homologue or fragment upon binding of an appropriate transcription factor to said promoter(s),
- 35 d) introducing said vector into said cell from

- 60 -

step a) comprising said transcription factor, and
e) identifying the phenotype of said cell
compared to wild type.

5 4. A method according to any of claims 1 to 3
wherein said DNA library, homologue or fragment is
cloned in a sense and an antisense direction relative
to said promoter.

10 5. A method according to any of claims 1 to 3
wherein said DNA library, homologue or fragment is
cloned between two promoters capable of producing
dsRNA from said DNA library, homologue or fragment
upon binding of said transcription factor to said
15 promoters.

 6. A method according to any of claims 1 to 5
wherein said cell is adapted to express said
transcription factor.

20 7. A method according to any of claims 1 to 6
wherein said DNA library, homologue or fragment is
constructed in a suitable vector which comprises a
sequence of nucleotides encoding said transcription
25 factor operably linked to a suitable promoter.

 8. A method according to any of claims 1 to 6
wherein said transcription factor is encoded by a
further vector independent of the vector including
30 said DNA library, DNA homologue or fragment and which
sequence encoding said transcription factor is
operably linked to a suitable promoter.

 9. A method according to claim 7 or 8
35 wherein said transcription factor comprises any of T7,

T3 or SP6 polymerase.

10. A method according to claim 7 or 8 wherein
said suitable promoter comprises any of let 858,
5 SERCA, UL6, myo-2 or myo-3.

11. A method according to any of claims 7 to 10,
wherein said suitable vector or said further vector
comprises a selectable marker.

10 12. A method according to claim 11 wherein said
selectable marker comprises a nucleotide sequence
capable of inhibiting or preventing expression of a
gene in said cell and which gene is responsible for
15 conferring a known phenotype.

13. A method according to claim 12 wherein said
nucleotide sequence comprises a sequence which is a
part of or identical to said gene conferring said
20 phenotype, and which nucleotide sequence is itself
oriented relative to a suitable promoter(s) capable of
initiating transcription of double stranded RNA upon
binding of an appropriate transcription factor to said
promoters.

25 14. A method according to claim 12 wherein said
nucleotide sequence is a part of or identical to said
gene sequence conferring said phenotype, and which
nucleotide sequence is such as to permit integration
30 of said suitable or further vector by homologous
recombination in the genome of said cell and following
said integration said nucleotide sequence is capable
of inhibiting expression of said gene sequence
conferring said phenotype.

35

- 62 -

15. A method according to claim 14 wherein said nucleotide sequence comprises stop codons sufficient to prevent translation of said nucleotide sequence following its integration into said genome.

5

16. A method according to any preceding claim wherein said cell is a microorganism suitable for feeding to, transforming or infecting an organism.

10

17. A method according to any of claims 1 to 14 wherein said cell is contained in an organism or an embryo thereof.

15

18. A method according to any of claims 1 to 17 wherein said promoters are T7 promoters.

20

19. A method according to any of claims 12 to 18 wherein said known gene sequence comprises a sup-35 gene or a fragment thereof which is selectable by identifying offspring growing at a temperature above 25°C following introduction of said vector in the genome of a pha-1 etl23ts mutant *C. elegans* worm.

25

20. A method according to any of claims 1 to 19 wherein said cell or organism is contacted with a specified compound for screening for a desired phenotype, such as resistance or sensitivity to said compound when compared to the wild type cell or organism.

30

21. A method according to any preceding claim wherein said transcription factor is inducible.

35

22. A method according to claim 16 wherein said microorganism is an *E. coli* strain which is an RNAase

III and preferably an RNAase negative strain.

23. A method according to any of claims 17
wherein said organism is selected from the class
5 nematoda.

24. A method according to claim 23 wherein said
organism is *Caenorhabditis elegans*.

10 25. A method of generating a transgenic non-
human organism comprising an exogenous transcription
factor and a transgene comprising a promoter operably
linked to DNA fragment which is expressed upon binding
of said transcription factor thereto, the method
15 comprising

a) providing a first transgenic organism
comprising a first construct incorporating DNA
encoding an exogenous transcription factor and a
second transgenic organism comprising a second
20 construct including at least one promoter operably
linked to a desired DNA sequence which is expressed
upon binding of the transcription factor of said first
transgenic organism thereto,

b) crossing said first and second transgenic
25 organisms and selecting offspring expressing said
desired DNA sequence.

26. A method according to claim 25 wherein said
first and second transgenic organisms are generated by
30 transforming said first and second constructs into
respective microorganisms for subsequent feeding to
the respective organism.

27. A method according to claim 25 or 26 wherein
35 said second construct comprises said desired DNA

sequence in an orientation relative to said promoter so as to be capable of initiating transcription of said DNA to dsRNA upon binding of said transcription factor thereto.

5

28. A method according to claim 27 wherein said second construct comprises two promoters flanking said desired DNA sequence which promoters can initiate transcription of said DNA sequence to dsRNA upon
10 binding of said transcription factor to said promoters.

29. A method according to claim 27 wherein said DNA sequence is provided in a sense and an antisense
15 orientation relative to said promoter so as to produce dsRNA upon binding of the transcription factor to the promoter.

30. A method according to any of claims 25 to 29
20 wherein said second transgenic organism further comprises a reporter gene operably linked to a promoter which is capable of initiating transcription of said reporter upon binding of said transcription factor thereto.

25

31. A method according to any of claims 25 to 30 wherein said transcription factor comprises a polymerase.

30

32. A method according to claim 31 wherein said polymerase comprises any of T7, T3 or SP6 polymerase.

33. A method according to any of claims 25 to 31 wherein said promoters comprises any of T7, T3 or SP6
35 promoters.

- 65 -

34. A method according to claim 30 wherein said reporter gene comprises any of those sequence encoding Luciferase, Green Fluorescent protein, β galactosidase or β -lactamase.

5

35. A method according to any of claims 25 to 30 wherein said organism is of the species nematoda.

36. A method according to claim 35 wherein said nematoda species is *C. elegans*.

10

37. A transgenic non-human multicellular organism obtainable according to the methods of any one of claims 25 to 34.

15

38. A method of validating clones identified in yeast two hybrid vector experiments which method comprises

- a) providing a construct including the DNA encoding the protein identified in the two hybrid vector experiment, which construct is such that said DNA is orientated relative to a promoter(s) that is capable of initiating transcription of said DNA to double stranded RNA upon binding of an appropriate transcription factor to said promoter(s),
- b) transforming a cell comprising said transcription factor with said construct, and
- c) identifying a phenotypic change in said cell or organism when compared to a wild type.

20

25

30

39. A method according to claim 38 wherein said DNA sequence is provided between two promoters capable of initiating transcription of the DNA sequence to dsRNA upon binding of the transcription factor to

35

- 66 -

said promoters.

40. A method according to claim 38 wherein said
DNA is provided in a sense and an antisense
5 orientation relative to said promoter such that
binding of the transcription factor to said promoter
initiates transcription of dsRNA from said DNA.

41. A method according to any of claims 38 to 40
10 wherein said transcription factor is inducible in said
cell.

42. A method according to any of claims 38 to 41
wherein said promoter is a phage polymerase promoter
15 and said transcription factor is a RNA polymerase.

43. A method according to claim 42 wherein said
polymerase is any of T7 RNA polymerase, T3 RNA
polymerase or SP6 RNA polymerase.
20

44. A method according to claim 43 wherein said
promoters comprise any of T7, T3 or SP6 promoter

45. A method according to any of claims 38 to 44
25 wherein said construct is such that it may be used in
yeast two hybrid experiments.

46. A method according to any of claims 38 to 45
wherein said cell is an *E. coli* cell.
30

47. A method according to any of claims 38 to 45
wherein said cell is part of an organism or an embryo
thereof.

35 48. A method according to claim 47 wherein said

- 67 -

organism is of the species nematoda and preferably *C. elegans*.

- 5 49. Plasmid pGN1 as illustrated in Figure 1.
50. Plasmid pGN100 as illustrated in Figure 2.
51. The yeast two hybrid vector plasmid
10 illustrated in any of Figures 4, 15 or 16 (Seq ID Nos 8
and 9).
52. A plasmid as illustrated in Figure 7.
53. A plasmid as illustrated in Figure 8.
- 15 54. A method of alleviating infestation of plant
pests, which method comprises
- a) identifying a DNA sequence from said pest
20 which is critical for its survival, growth,
proliferation,
- b) cloning said sequence from step a) or a
fragment thereof in a suitable vector in an
orientation relative to promoter(s) such that said
promoter(s) is capable of initiating transcription of
25 said DNA sequence to RNA or dsRNA upon binding of an
appropriate transcription factor to said promoter(s),
and
- c) introducing said vector into the plant.
- 30 55. A method according to claim 54 wherein said
DNA sequence is provided between two promoters such
that binding of the transcription factor to the
promoters results in transcription of the DNA to
dsRNA.

35

- 68 -

56. A method according to claim 54 wherein said DNA sequence is provided in a sense and an antisense orientation relative to said promoter such that binding of the transcription factor to the promoter results in transcription of the DNA to dsRNA.

57. A method according to claim 54 wherein said pest is a nematode worm.

58. A method according to claim 57 wherein said nematode comprises any of *Tylenchulus* ssp., *Radopholus* ssp., *Rhadinaphelenchus* ssp., *Heterodera* ssp., *Rotylenchulus* ssp., *Pratylenchus* ssp., *Belonolaimus* ssp., *Canjanus* ssp., *Meloidogyne* ssp., *Globodera* ssp., *Nacobbus* ssp., *Ditylenchus* ssp., *Aphelenchoides* ssp., *Hirschmenniella* ssp., *Anguina* ssp., *Hoplolaimus* ssp., *Heliotylenchus* ssp., *Cricone mellasp.*, *Xiphinemasp.*, *Longidorus* ssp., *Trichodorus* ssp., *Paratrichodorus* ssp., *Aphelenchs* ssp.

59. A method according to claim 55 wherein said DNA sequence or fragment thereof is cloned between two tissue, preferably root specific promoters.

60. An expression vector for use in a method according to any proceeding claim comprising a promoter or promoters oriented relative to a DNA sequence such that they are capable of initiating transcription of said DNA sequence to double stranded RNA upon binding of an appropriate transcription factor to said promoter or promoters.

61. An expression vector according to claim 60 comprising two identical promoters flanking the DNA sequence.

- 69 -

62. An expression vector according to claim 60 comprising said DNA sequence in a sense and an antisense orientation relative to said promoter.

5 63. An expression vector according to any of claims 60 to 62 which further comprises a nucleotide sequence encoding a selectable marker.

10 64. An expression vector according to claim 63 wherein said nucleotide sequence encoding said selectable marker is orientated relative to the promoter(s) such that transcription of the nucleotide sequence to double stranded RNA occurs upon binding of an appropriate transcription factor to said
15 promoter(s).

20 65. An expression vector according to claim 64 wherein said nucleotide sequence encoding the selectable marker is provided between the identical promoters capable of initiating transcription of the
25 nucleotide sequence to dsRNA upon binding of the transcription factor to the promoters.

30 66. An expression vector according to claim 64 wherein said nucleotide sequence encoding the selectable marker is provided in a sense and an antisense orientation relative to the promoter such that transcription of the nucleotide sequence to dsRNA upon binding of the transcription factor to said
35 promoter occurs.

67. An expression vector according to claim 63 or 64 wherein said selectable marker comprises a nucleotide sequence encoding sup-35, for introduction
35 into *C. elegans* having a pha-1 mutation.

- 70 -

68. An expression vector for expressing an appropriate transcription factor for use in a method according to any of claims 1 to 48 and 54 to 59 which vector comprises a sequence of nucleotides encoding
5 said transcription factor operably linked to suitable expression control sequences.

69. An expression vector according to claim 68 wherein said expression control sequences include
10 promoters which are inducible, constitutive, general or tissue specific promoters, or combinations thereof.

70. An expression vector according to any of claims 68 to 69 wherein said transcription factor
15 comprises a phage polymerase, and preferably T7 RNA polymerase.

71. An organism or cell transformed or transfected with a plasmid according to any of claims
20 49 to 53 or an expression vector according to any of claims 60 to 70.

72. An organism according to claim 71, which is of the species nematoda and preferentially *C. elegans*.
25

73. A method of introducing dsRNA or DNA capable of producing dsRNA into an organism which method comprises feeding said organism with a suitable microorganism comprising an expression vector
30 according to any of claims 60 to 67 or feeding said organism directly with an expression vector according to any of claims 60 to 67.

74. A method according to claim 73 wherein said
35 microorganism or said organism is adapted to express

- 71 -

said transcription factor.

75. A method according to claim 74 wherein
either said microorganism or said organism comprises
5 an expressing vector according to any of claims 66 to
69.

76. A method according to any of claims 73 to 75
wherein said organism is *C. elegans* and said
10 microorganism is *E. coli*.

77. A method according to claim 76 wherein said
E. coli strain is an RNAaseIII negative strain.

15 78. A method according to any of claims 73 to 75
wherein said organism is a *C. elegans* nuc-1 mutant.

79. A selection system for identifying
transformation of a cell or organism with a vector
20 according to claims 60 to 63 which system comprises a
vector according to claims 60 to 63 and said
selectable marker comprises a nucleotide sequence
capable of inhibiting or preventing expression of a
gene in said cell or organism which gene is
25 responsible for conferring a known phenotype.

80. A selection system according to claim 79
wherein said nucleotide sequence comprises a sequence
which is a part of or identical to said gene
30 conferring said known phenotype, and which nucleotide
sequence is itself located between two identical
promoters capable of initiating transcription of
double stranded RNA upon binding of an appropriate
transcription factor to said promoters.

35

81. A selection system according to claim 79 wherein said nucleotide sequence is a part of or identical to said gene sequence which confers a known phenotype on said cell or organism, and which
5 nucleotide sequence permits integration of said vector by homologous recombination in the chromosome of said cell or organism and following said integration said sequence inhibits expression of said gene sequence conferring said known phenotype.

10 82. A selection system according to claim 81 wherein said nucleotide sequence comprises stop codons sufficient to prevent translation of said nucleotide sequence following integration into said genome.

15 83. A selection system according to claim 79 wherein said known gene sequence comprises a sup-35 gene or a fragment thereof which is selectable by identifying offspring growing at a temperature above
20 25°C following introduction of said vector in a pha-1 et123ts mutant *C. elegans* worm.

84. A method according to claim 74 wherein said transcription factor is T7 RNA polymerase.

25 85. A method of assigning function to a DNA sequence of a multicellular organism which method comprises:

a) providing:

30 (i) a first construct comprising said DNA sequence cloned in a sense direction under the regulation of a suitable promoter

(ii) a second construct comprising said DNA sequence cloned in an anti-sense direction under the
35 regulation of the same promoter as in step (i),

in a multicellular organism which is adapted to initiate transcription of said DNA fragment from said promoter,

- 5 b) identifying the phenotype of said multicellular organism compared to wild type.

86. A method of assigning function to a DNA sequence of a multicellular organism which method comprises:

- 10 a) providing
 i) a construct comprising said DNA fragment cloned between two promoters capable of promoting transcription in said multicellular organism,
 in a multicellular organism capable of
15 initiating transcription from said promoter;
 b) identifying the phenotype of said multicellular organism compared to wild type.

87. A method for expressing a gene, cDNA or
20 DNA fragment in *C. elegans* which method comprises :
 i) providing a transgenic *C. elegans* expressing an exogenous transcription factor,
 ii) cloning said gene or cDNA or DNA fragment into an appropriate vector operably linked to a
25 promoter capable of initiating transcription of said gene, cDNA or DNA fragment upon binding of said exogenous transcription factor thereto,
 iii) introducing said vector into said transgenic *C. elegans* from step i).

30

88. A method according to claim 87 wherein said exogenous transcription factor is a phage polymerase.

35

89. A method according to claim 88 wherein

- 74 -

said phage polymerase is T3, SP6 or T7 RNA polymerase.

90. A method according to any of claims 87 to
89 wherein said promoter is any of T3, T7 or SP6
5 promoter.

1/21

FIG. 1.

pGN1

gagtgcaccatatcgcggtgtgaaataaccgcacagatgcgtaaggagaaataaccgcacagcgaaattgtaaacgttaatatatt
tgtaaaaattcggttaaatattgttaaatcagctcatttttaaccaataggccgaaatcggcataaatccctataaatcaaaagaat
agaccgagatagggttgagtggttccagtttgaacaagagtgccactattaaagaacgtggactccaacgtcaaaaggcgga
aaaccgtctatcaggggcgatggccactacgtgaaccatcacccaaatcaagtttttgcggtcgaggtgccgtaaagctctaatt
cggaaccctaaaggagacccccgatattagagcttgacggggaaagccggcgaaacgtggcgagaagggaagggaagaaag
cgaaaggagcgggcgctaggcgctggcgaagtgtagcgggtcacgctgcgctaaccaccacaccgcgcgcttaattgcgc
cgctacaggggcgctccattcgccattcagggtcgcgcaactgttgggaaggcgatcggtcggggcctcttcgctattacgcca
gctggcgaaaagggggatgtgctgcaaggcgattaaagttgggtaacgccagggttttccagtcacgacgttgtaaaccgacgg
ccagtgaaattgtaatacgactcactataggggcgaattcgagctcggtaccggggatcctctagagtcgaaagcctcgcacctat
agttagctcgtattacagcttgagttattctatagtgacctaataatagcttgccgtaatacatggtcatagctgtttcctgtgtgaaattgtt
atccgctcacaaattccacacaacatacagagccgggaagcataaagtgtaaagcctgggggtgctaataagtgagctaaactcacatt
aattgcgttgcgctcactgccccgtttccagtcgggaacacctgtcgtgccagctgcattaatgaatcgccaacgcgcggggag
aggcggtttgcgattggggcgtcttcgcttctcgtcactgactcgctcgctcggtcggttcggctcgggcgagcggtatcag
ctcactcaaaaggcggttaatacggttatccacagaatcaggggataaacgcaggaaagaaatgtgagcaaaaggccagcaaaa
ggccagggaaccgtaaaaaggcgcggttgctggcgcttttcgataggctccgccccctgacgagcatcaaaaaatcgacgct
caagtcagaggtggcgaaaccgcagaggactataaagataccaggcggttccccctggaagctccctcgtgcgtctcctgttc
cgacctgcccgttaccggataacctgtccgcctttctccctcgggaagcgtggcgctttctcatagctcacgctgtaggtatctca
gttcgggtgtaggtcggtcgctccaaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggttaacta
tcgtcttgagtgccaacccggtgaagacacgacttatcgccactggcagcagccactggaacaaggattagcagagcgaggtatgt
aggcggtgctacagagttctgaaagtgtggcctaactacggctacactagaaggacagattttggtatctgcgctctgctgaagc
cagttaccttcggaaaaaagagttgtagctcttgatccggcaaaacaaaccaccgctggtagcggttggtttttttgttgcaagcagc
agattacgcgcagaaaaaaaggatctcaagaagatccttgatcttttctacgggggtctgacgctcagtggaacgaaaactcacgt
taaggggattttggtcatgagattcaaaaaggatcttcacctagatccttttaaatataaaatgaagtttaaatcaatctaaagtatat
atgagtaaacctggctgcagcttaccatgtcttaatacagtgaggcacctatctcagcgatctgtctatttcgttcacatagttgctt
gactccccgtcgtgtagataactacgatacgggaggggttaccatctggccccagtgctgcaatgataccgcgagaccacgct

2/21

FIG. 1. (CONTINUED)

caccggctccagattatcagcaataaaccagccagccggaaggccgagcgcagaagtggctcctgcaactttatccgcctcc
 atccagtctattaattgttgcgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttggcattgctacaggca
 tcgtggtgtcacgctcgtcgttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccattgtgtg
 caaaaaagcggtagctccttcggtcctccgatcgttgcagaagtaagtggccgcagtggttatcactcatggttatggcagcact
 gcataattcttactgtcatgccatccgtaagatgctttctgtgactggtgagtactcaaccaagtcattctgagaataccgcgccc
 ggcgaccgagttgctctgcccggcgtaatacgggataatagtgtatgacatagcagaactttaaaagtgtcatcatggaaaa
 cgttcttcggggcgaaaacttcaaggatcttaccgctgttgagatccagttcgtatgtaaccactcgtgcacccaactgatcttca
 gcatctttactttaccagcgttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataaggcgacacg
 gaaatgtgaataactcatacttctcttttcaatattatgaagcatttatcagggttatgtctcatgagcggatacatattgaatgtat
 ttagaaaaataaacaataagggttccgcgcacattccccgaaaagtgccacctgacgtctaagaaaccattattatcatgacatt
 aacctataaaaaataggcgtatcacgaggcccttctgtctcgcgcttgcgtgatgacggtgaaaaccttgacacatgcagctcc
 cggagacgggtcacagcttctgtgaagcggatgccgggagcagacaagcccgtcagggcgctcagcgggtgttggcgggt
 gtcggggctggcttaactatcgggcatcagagcagattgtactga

FIG. 2.

PGN100

ctagcatgaacacgattaacatcgctaagaacgacttctctgacatcgaactggctgctatcccgttcaacactctggctgaccatt
 acggtgagcgtttagctcggaacagttggcccttgagcatgagttacgagatgggtgaagcacgcttccgcaagatgtttga
 gcgtcaactaaagctgggtgaggttgcggataacgctgcgcgaagcctctcatcactacctactccctaagatgattgcacgc
 atcaacgactggttgaggaagtgaagctaagcgcggcaagcgcggacagccttccagttcctgcaagaaatcaagccgga
 agccgtagcgtacatcaccanaagaccactctggcttgcctaaccagtgtgacaatacaaccgttcaggctgtagcaagcgca
 atcggctcggccattgaggacgaggctcgtcttcggtcgtatccgtgacctgaagctaagcacttcaagaaaaacgttgaggaa
 caactcaacaagcgcgtagggcacgtctacaagaaagcatttatgcaagttgtcagggtgacatgctcttaagggtctactcg
 gtggcgaggcgtggtcttcgtggcataaggaaagactctattcatgtaggagtacgctgcatcgagatgctcattgagtcaaccgg
 aatgggttagcttacaccgccaaaatgctggcgtagtaggtcaagactctgagactatcgaactcgcacctgaatacgtgaggct
 atcgaaccctgacaggtgcgtggtggtcatctctcgatgttccaaccttgcgtagttcctcctaagccgtggactggcattac
 tgggtggtggttattgggctaaccgtcgtcgtcctctggcgctggtgctgactcacagtaagaaagcactgatgcgtacgaagac
 gttacatgcctgaggtgtacaaagcgattaacattgcgcaaaacaccgcatggaaaatcaacaagaaagtcctagcggctgcc
 aacgtaatcaccaagtggaaagcattgtccggtcaggacatccctgcgattgagcgtgaagaactcccgatgaaaccggaaga
 catcgacatgaatcctgaggctctcaccgcgtggaaacgtgctgcgctgctgttaccgcaaggacagggtcgcgaagtctc
 gccgtatcagccttgagttcatgcttgagcaagccaataagtttctaaccataaggccatctggttcccttacaacatggactggc
 gcggtcgtgtttacgcccgtgtcaatgttcaaccgcaaggtaacgatatgaccaaaggactgcttaccgtggcgaaaggtaaac
 caatcggttaagggaaggttactactggctgaaaatccacgggtgcaactgtgcgggtgtcgataaggttccgttccctgagcgc
 caagttcattgaggaaaaccacgagaacatcatggttgcgctaagtctccactggagaacacttgggtgggctgagcaagattct
 ccgttctgttcccttgcgttctgctttagtacgctgggggtacagcaccacggcctgagctataactgttcccttccgctggcgtttg
 acgggtcttctgcttggcatccagcacttctccgcgatgctccgagatgaggttaggtggtcgcgcgggttaactgttcttagtgag
 accgttcaggacatctacgggtattgttctaagaaagtcaacgagattctacaagcagacgcaatcaatgggaccgataacgaa
 gtagttaccgtgaccgatgagaacactgggtgaaatctctgagaaagtcaagctgggcactaaggcactgggtggtcaatggctg
 gctcacggtgttactcgcagtgtactaagcgttcagtcatgacgctggcttaccgggtcacaagagttcggcttccgtcaacaagt
 gctggaagataccattcagccagctattgattccggcaagggtccgatgttactcagccgaatcagcgtgctggatacatggct
 aagctgatttgggaatctgtgagcgtgacgggtgtagctgcggttgaagcaatgaactggcttaagtctgctgtaagctgctgg

3/21

FIG. 2. (CONTINUED 1)

ctgctgaggtcaagataagaagactggagagattcttcgaagcgttcgctgtgcattgggtaactcctgatggttccctgtgt
 ggcaggaaatacaagaagcctattcagacgcgcttgaaacctgatgttcctcggcagttccgcttacagcctaccattaacaccaac
 aaagatagcagagattgatgcacacaaacaggagctctggtatcgctcctaactttgtacacagccaagacggtagccaccttcgta
 agactgtagtgtgggcacacgagaagtaggaatcgaatctttgactgattcacgactccttcggtagccattccggctgacgct
 gcgaacctgttcaaagcagtcgcgaaactatggttgacacatatgagcttctgtatgtactggctgatttctacgaccagttcgct
 gaccagttgcacgagttcaattggacaaaatgccagcacttcggcctaaaggtaacttgaacctccgtgacatcttagagtcgg
 acttcgcttcgctgaacctggtattgatatctgagctccgcatcgccgctgtcatcagatcgccatctcgccgctgctcctg
 acttctaagtcgaattactctcaacatccctacatgctcttctccctgtgctccacccccctattttgttattatcaaaaaacttctc
 ttaatttcttgtttttagcttctttagctcaccttaacaatgaaattgtgtagattcaaaaatagaatttaattcgtaataaaaaagtcga
 aaaaaattgtgctccctcccccttaataataattctatcccaaatctacacaaatgttctgtgtacacttctatgttttttacttctga
 taaattttttgaaacatcatagaaaaaacgcacacaaaatacctatcatatgttacgtttcagttatgaccgcaattttatttcttc
 gcacgtctgggctctcatgacgtcaaatcaatcgtcatcgtgaaaaagtttgagtagttttggaattttcaatcaagtgaagttta
 tgaatttaatttctgcttttgggggttcccctattgtttgtcaagagtttcgaggacggcgttttcttgctaaaatcacaag
 tattgatgagcacgatgcaagaaagatcggaagaaggtttgggttgaggctcagtggaaggtgagtagaagttgataattgaa
 agtggagtagtgtctatgggggtttgccttaaatgacagaatacattcccaatataccaaacataactgttctactagtcggcgt
 acgggccccttcgtctcgcggttcgggtgatgacggtgaaaacctctgacacatgcagctcccggagacgggtcacagcttctgt
 gtaagcggatgcccgggagcagacaagcccgtcaggggcgtcagcgggtgtggcgggtgtcgggggtggttactatgc
 ggcatcagagcagattgtactgagagtgacacatagcgggtgtgaaataaccgcacagatgctgaaggagaaaaataccgcatca
 ggcgcccttaaggggcctcgtgatacgctattttataggttaattgtcatgataataatgggttcttagacgtcaggtggcacttttcg
 gggaaatgtgcgcggaacccctatttgttttttctaataacattcaaatatgtatccgctcatgagacaataacccgtgataaatgt
 tcaataatattgaaaaaggaagagtatgagtagttcaacattccgtgtcgccttattccctttttgcggttattgcttctgttttg
 ctaccacagaaacgctggtgaaagttaaagatgtgaagatcagttgggtgcacgagtggttacctgaactggatctcaaca
 gcggttaagatccttgagagtttcgccccgaagaacgtttccaatgatgagcacttttaagttctgtatgtggcgcggtattatc
 ccgtattgacgcccgggcaagagcaactcggctcggcgcatacactattctcagaatgacttggttgagtagtaccagtcacagaa
 aagcatctacggatggcatgacagtaagagaattatgacgtgtgccataaccatgagtgataaactgcggccaacttactct
 gacaacgatcggagaccgaaggagctaacccgtttttgcacaacatgggggatcatgtaactcgccttgatgttgggaacc
 ggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgctgtagcaatggcaacaacgttgcgcaaaactattaa
 ctggcgaaactactacttagcttccgggaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgtc
 ggcccttcggctggtgtgttattgtgataaatctggagcgggtgagcgtgggtctcgcggtatcatgacgactggggcca
 gatgtgaagccctcccgatcgtatgtatctacacgacggggagtcagggaactatggatgaacgaaatagacagatcgttag
 ataggtgcctcactgattaagcattggttaactgtcagaccaagttactcatatatacttttagattttaaaacttatttaattaaa
 aggatctaggtgaagatccttttgataatctcatgacaaaaatcccttaacgtgagtttctgtccactgagcgtcagaccccgtag
 aaaagatcaaaggatctcttgagatcctttttctgcgcgtaactctgctgttgcaaacaaaaaaaccacgctaccagcgggtg
 ttgtttgcggatcaagagctaccaactcttttcgaaggttaactggcttcagcagagcgcagataccaaatactgtcctttagt
 gtagccgtatgtagccaccactcaagaactctgtagcaccgctacatacctcgtctgtatctgttaccagtggctgctgc
 cagtggcgataagtcgtgttaccgggttgactcaagacgatagttaccggataaaggcgcagcggctcgggctgaacggggg
 gttcgtgcacacagcccagcttgagcgaacgacctacaccgaactgagatacctacagcgtgagcattgagaagcgcac
 gcttccggaaggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcagcaggaggagcttcca
 gggggaaacgcctggatctttatagtcctgtcgggttcgccaccttgacttgagcgtcgattttgtgatgtcgtcagggggg
 cggagcctatggaaaacgccgaacgcggccttttacggttctggttctggttctggttctcacatgttcttctcgtt
 atccctgattctgttgataaccgtataccgctttgagtgagctgataccgctcggcgcagccgaacgaccgagcgcagcga
 gtcagttagcggaggaagcgaagagcggcccaatacgaacacgcctctccccgcggttggccgattcaatgacgtgg
 cacgacaggtttccgactgaaaagcgggcagtgagcgaacgcaatattgtgagttagctcactcattaggcaccgccaggt
 ttacatttatgttccggctcgtatgtgtgtgaattgtgagcggataacaattcacacaggaaacagctatgaccatgattacg
 ccaagcttgcagcctgcaggtcgaacttagaggatcaagagcattgaatcagaatatggagaacggagcatgagcatttgcga
 agttttagatgactagaacaaagcgtgttggttccctctgagcccgcttcttatatacccgacttgcagccttacagaatgt

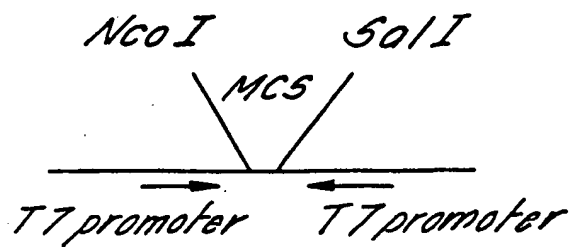
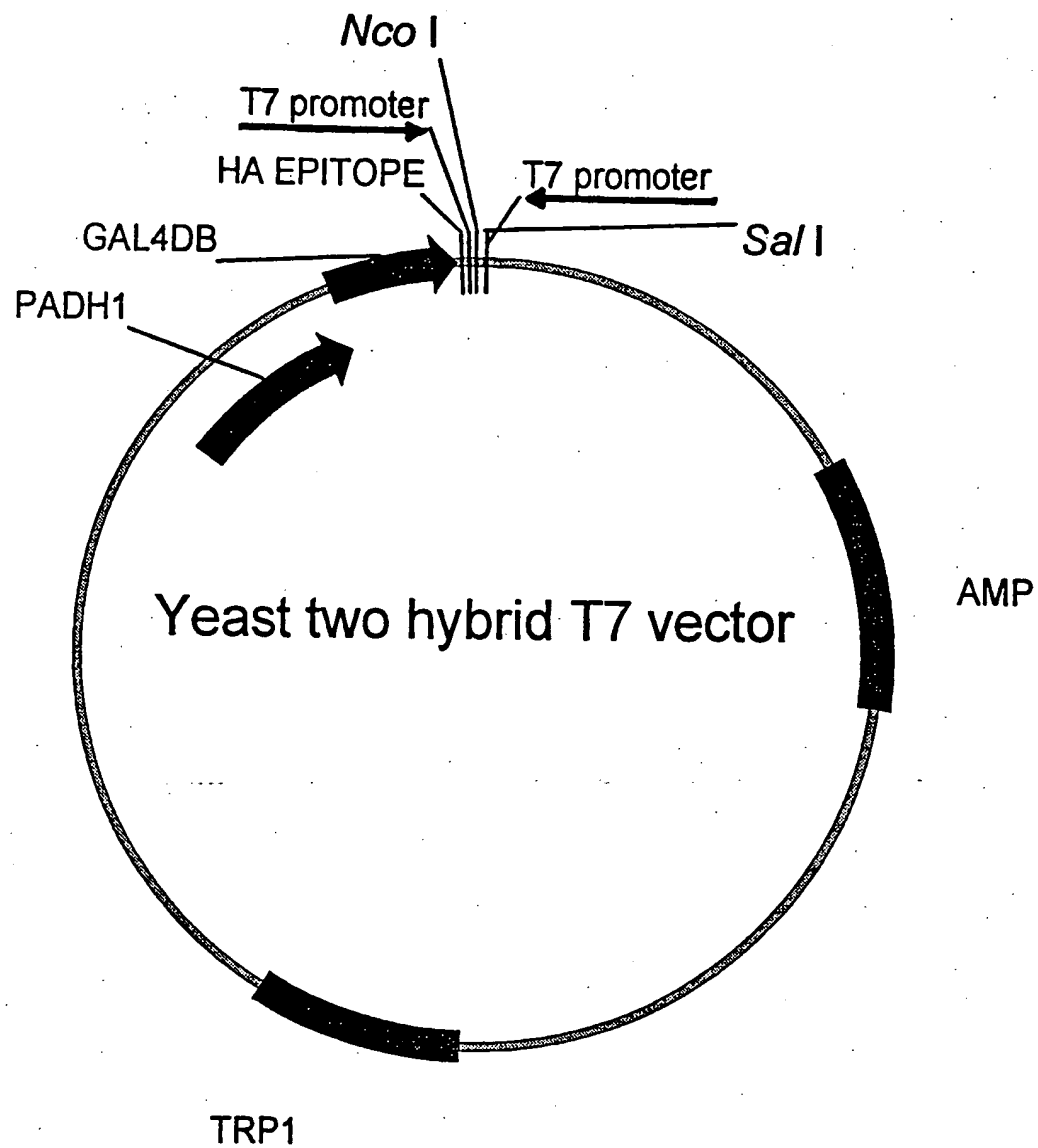
4/21

FIG. 2. (CONTINUED 2)

ctagaaggctctagatgcattcggttgaaaatactcccgggtgggtgcaaagagacgcagacggaaaatgtatctgggtctcttatt
gtgtacactacttttccatgtaccgaatgtgagtcgccctccttttgcaacaagcagctcgaatgttctagaaaaagggtggaaaata
gtataaataccgttgaaaataataccgaacaacatttgctctaattgtgaaattagaaatcttcaactataatcatctcactggatc
cccgggattggccaaaggacccaaagggtatgttcgaatgatactaacataacatagaacattttcaggaggacccttgg

6/21

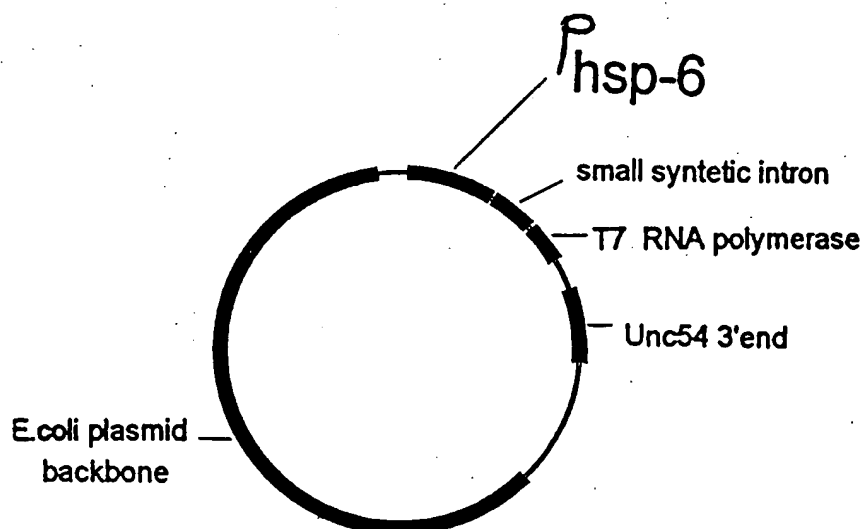
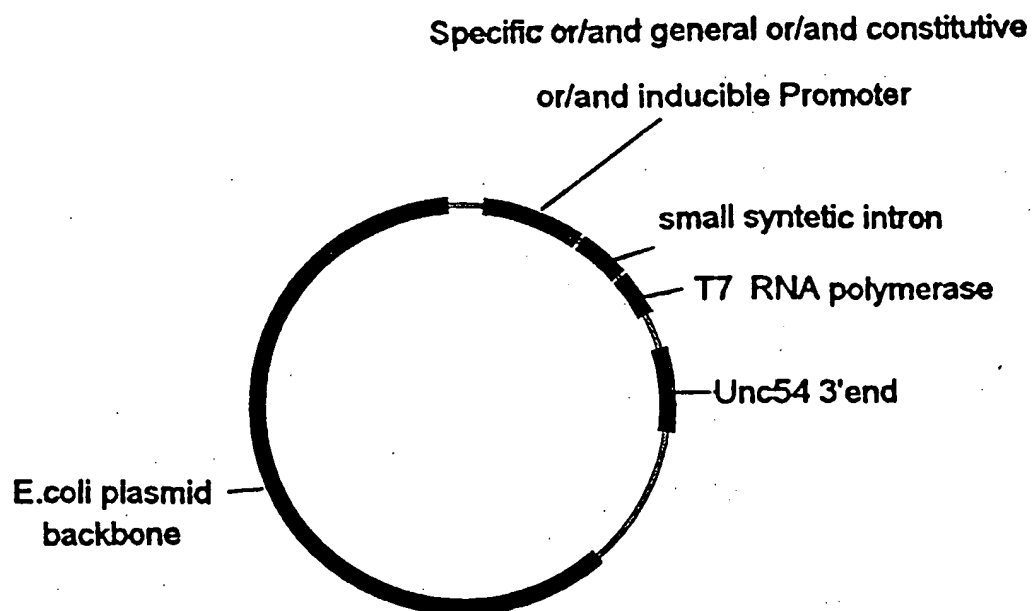
FIG. 4.



7/21

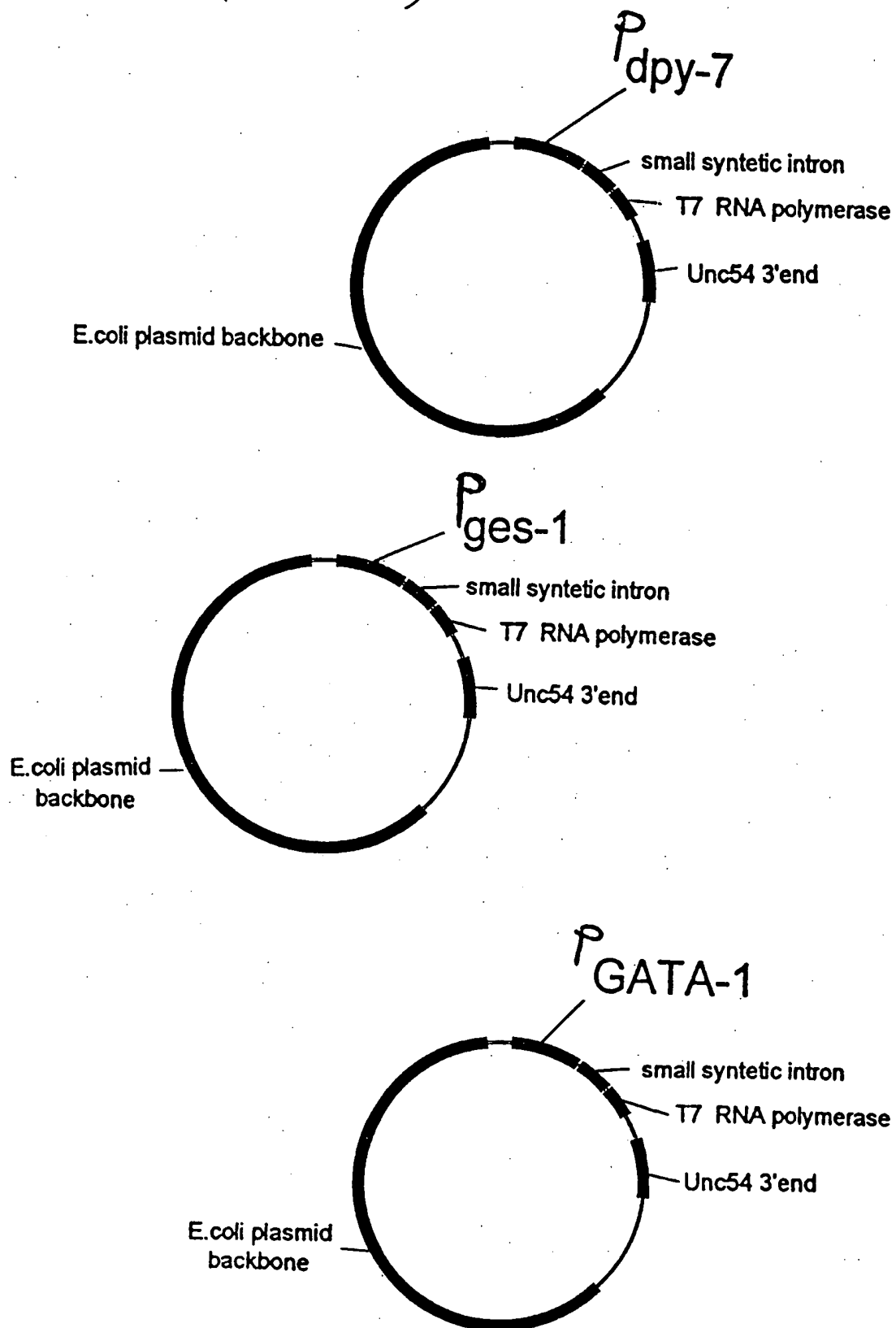
FIG. 5.

General description of the *C.elegans* T7 RNA polymerase
expression vector with 4 examples



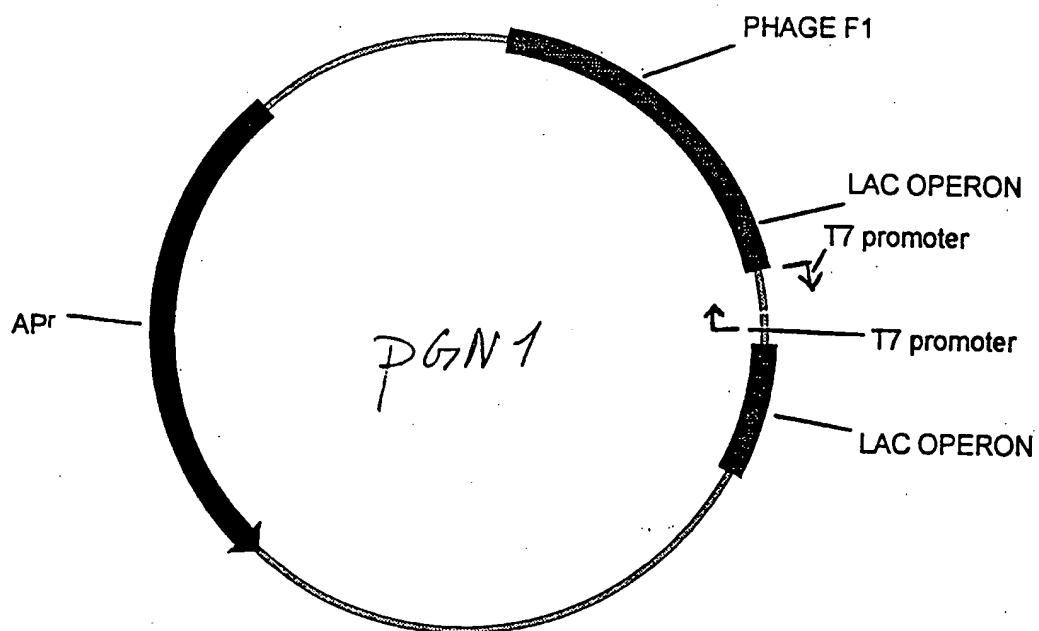
8/21

FIG. 5. (CONTINUED)



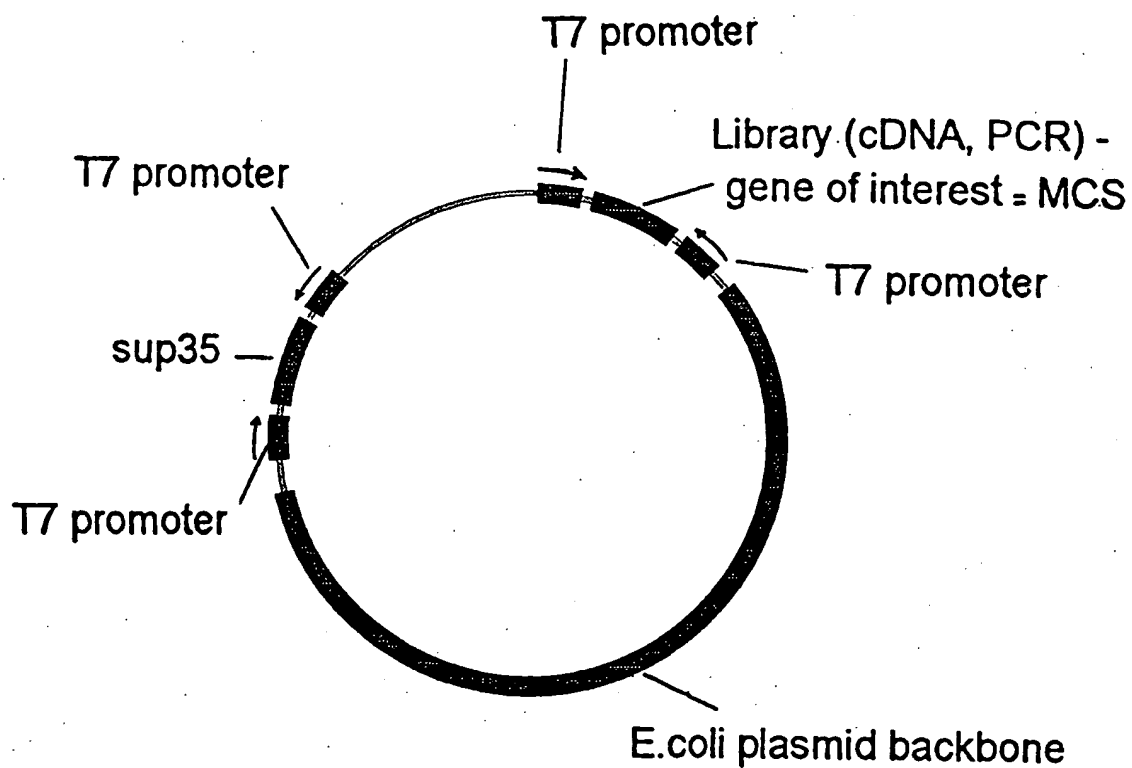
9/21

FIG. 6.



10/21

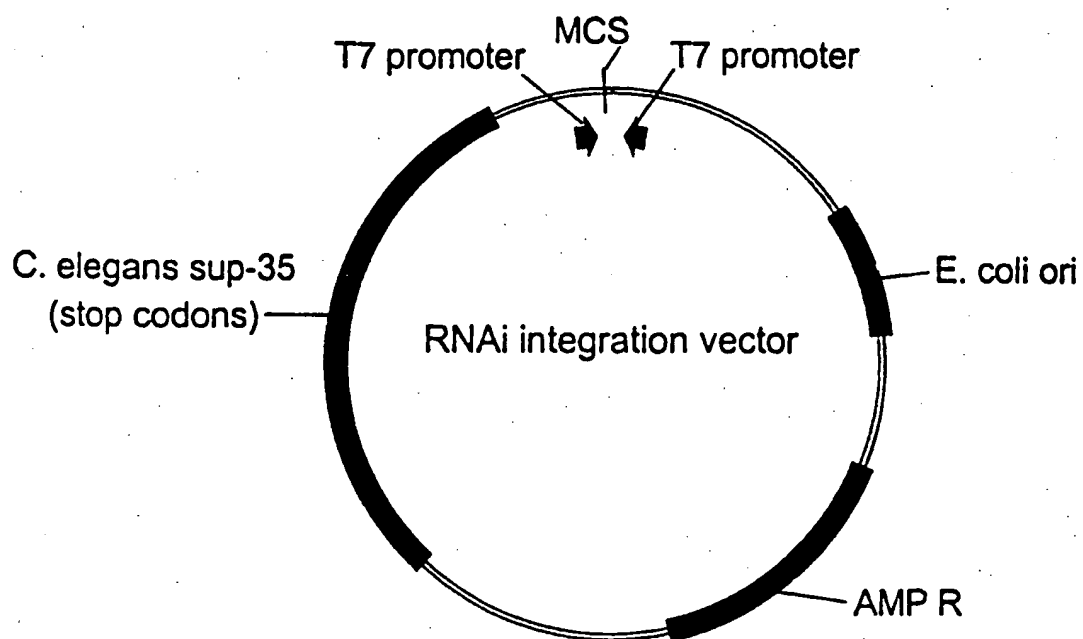
FIG. 7



*enhanced vector for RNAi, producing
sup35 dsRNA and dsRNA of the library,
gene of interest or PCR product.*

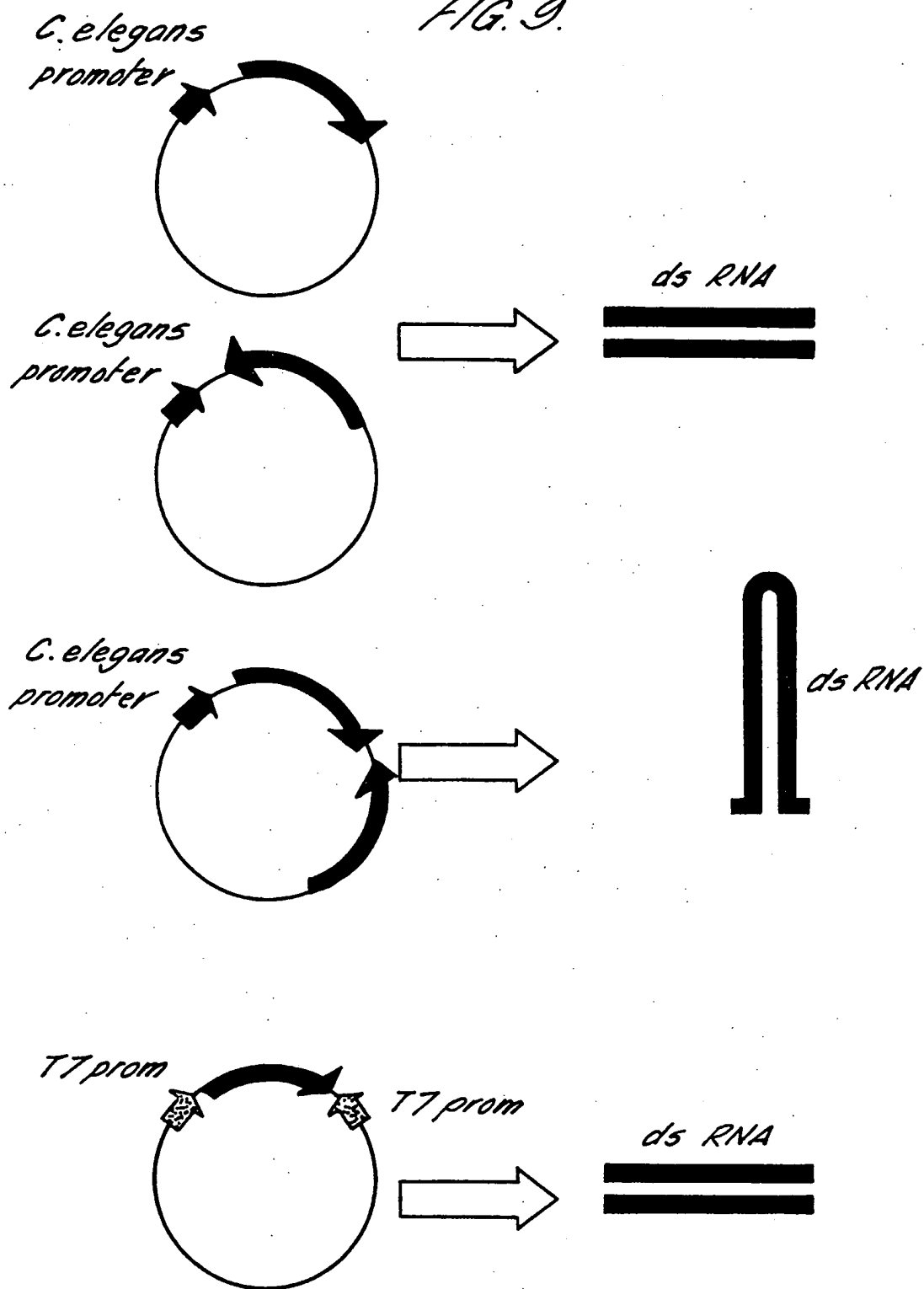
11/21

FIG. 8.



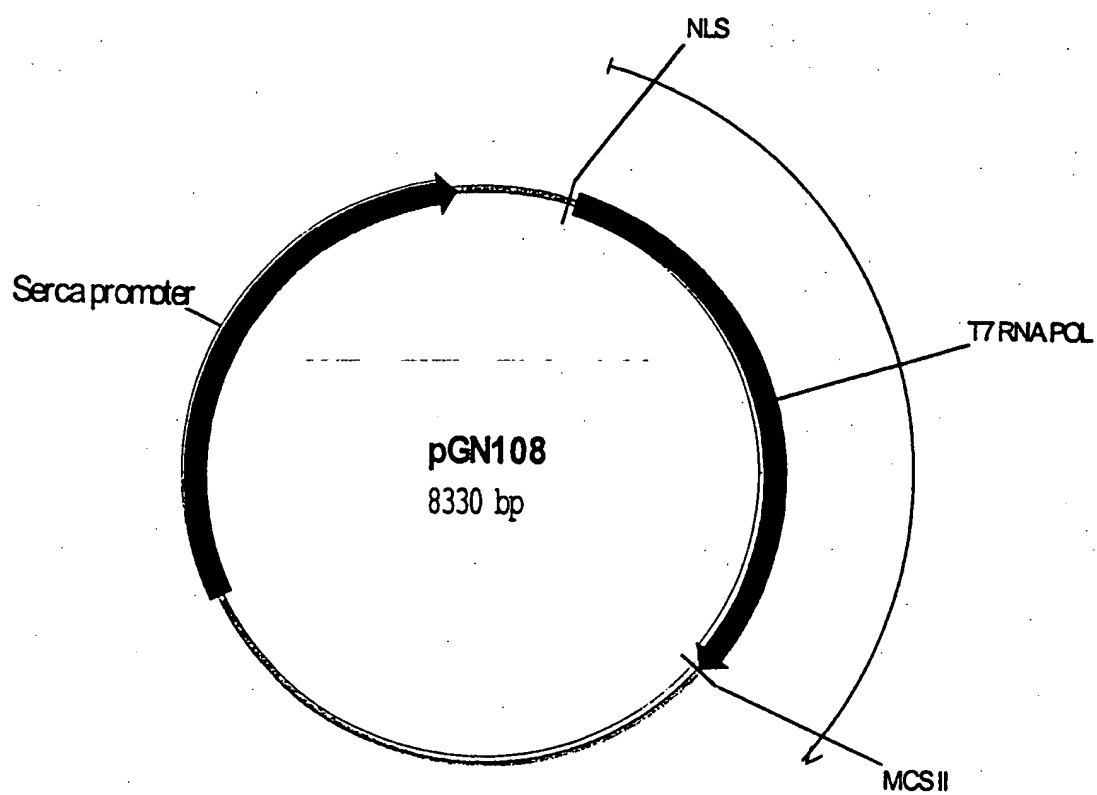
12/21

FIG. 9.



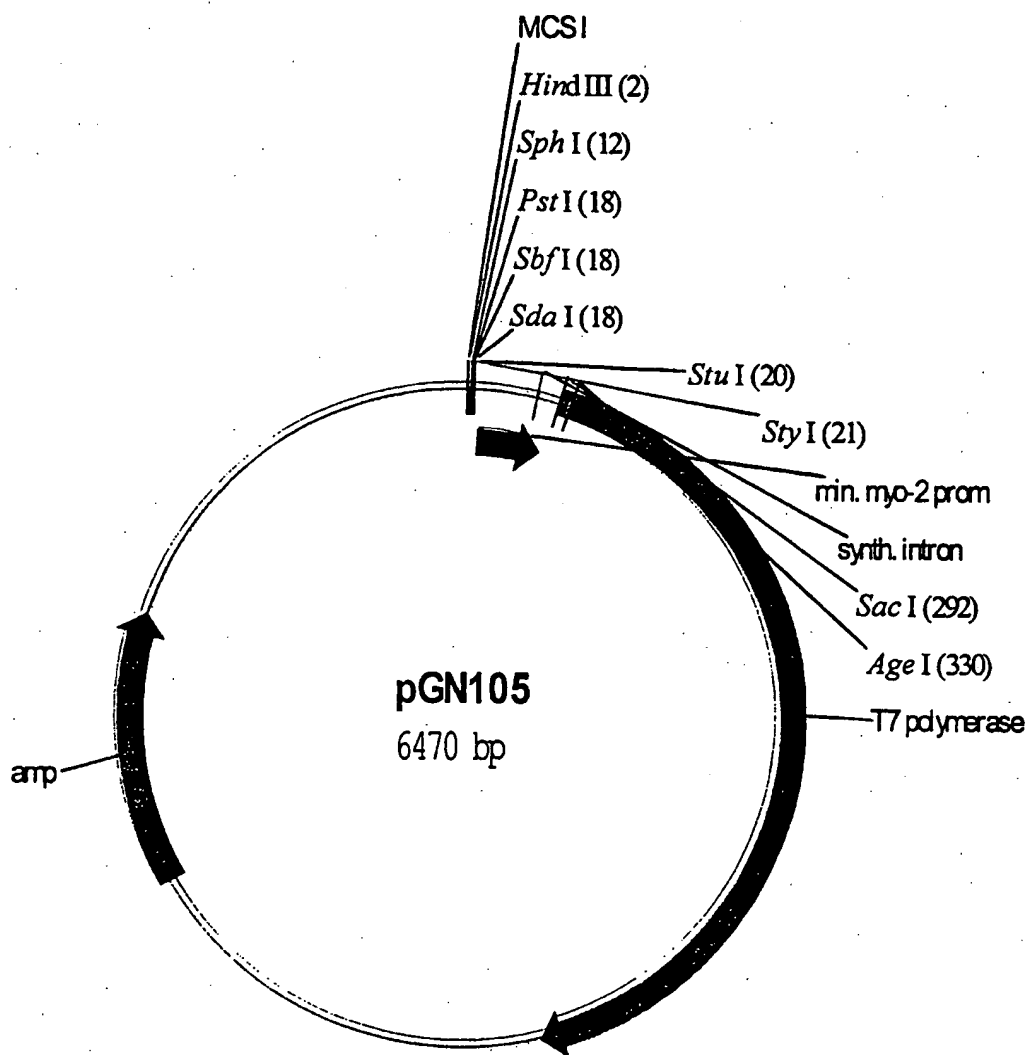
13/21

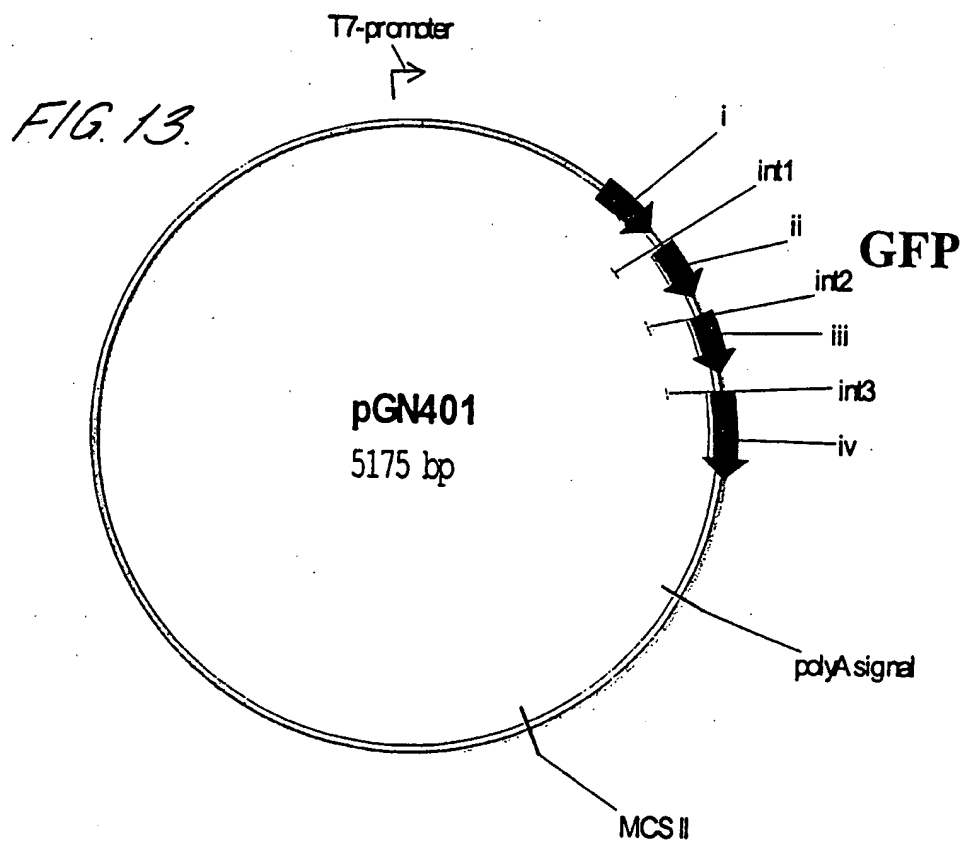
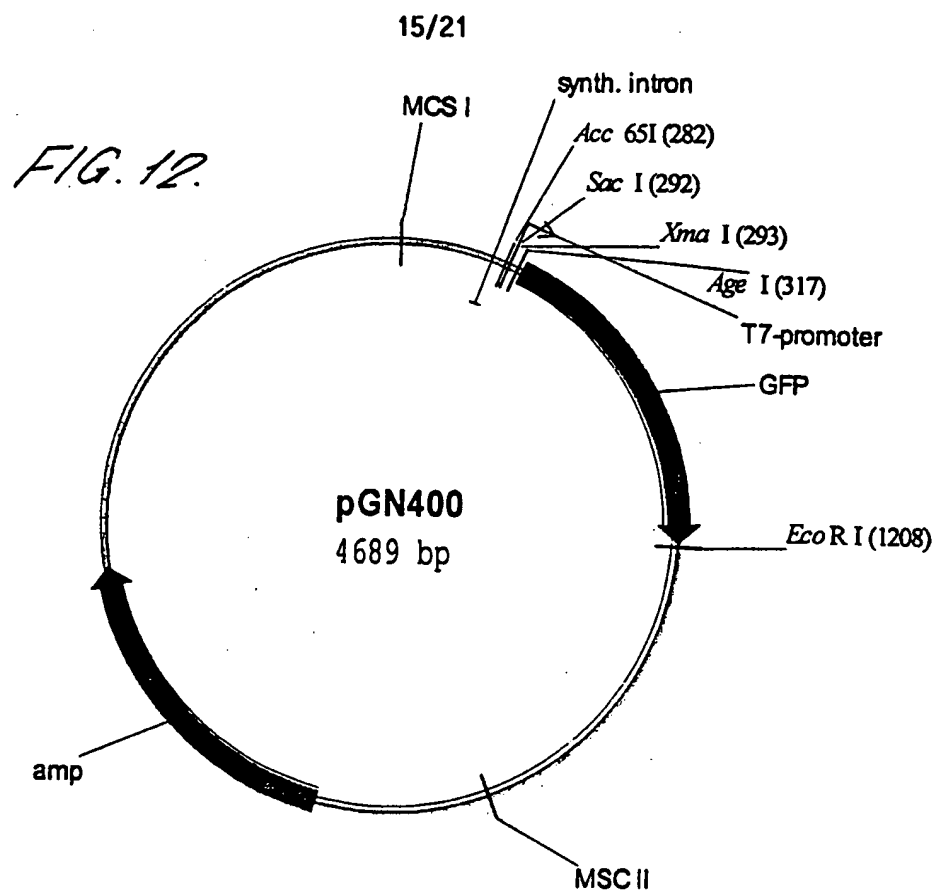
FIG. 10.



14/21

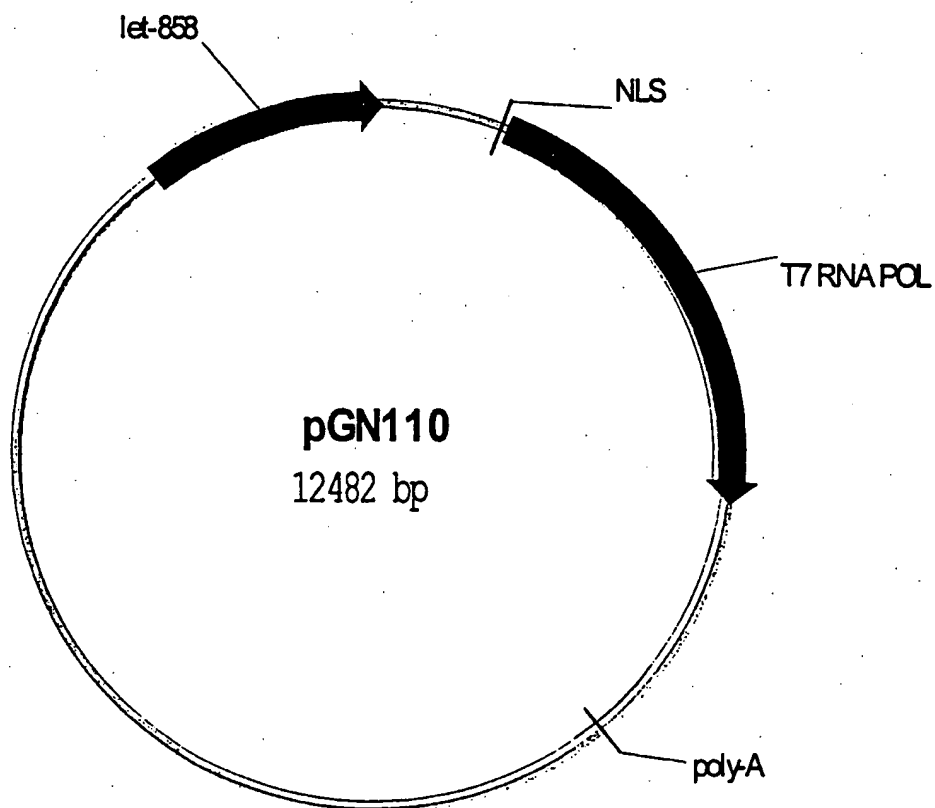
FIG. 11.



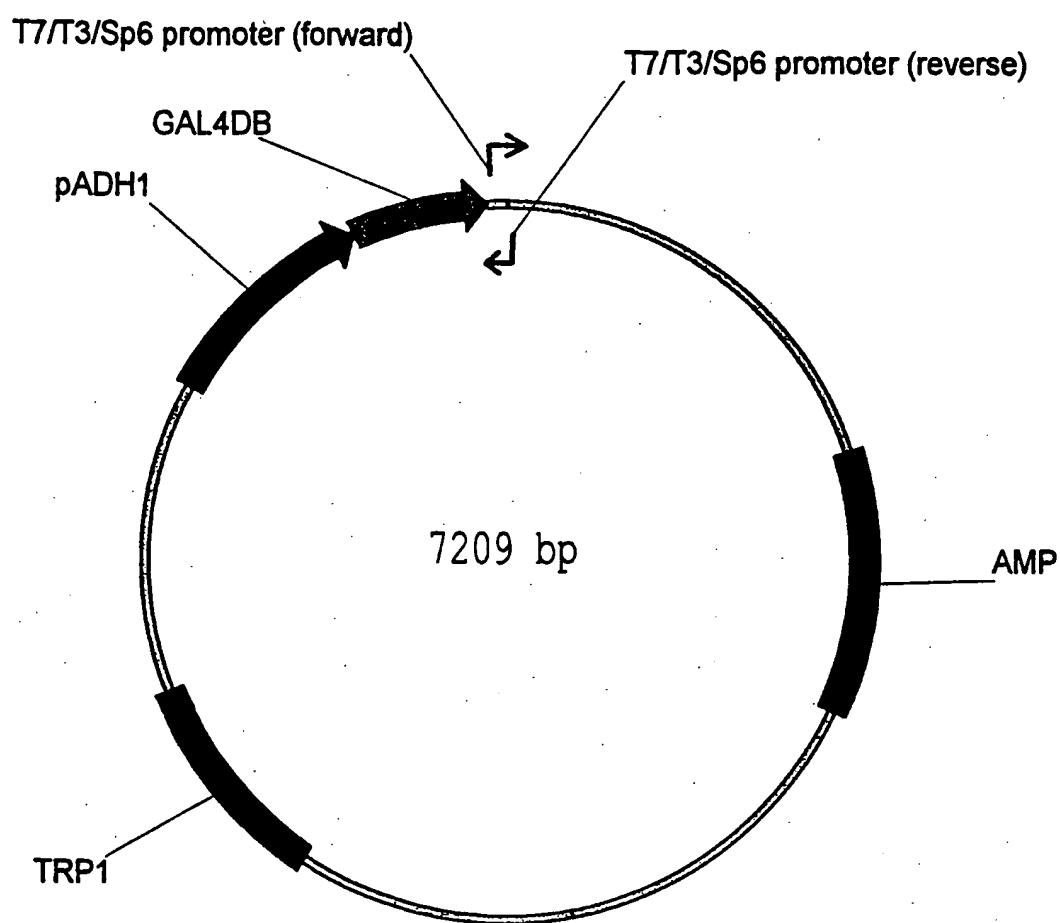


16/21

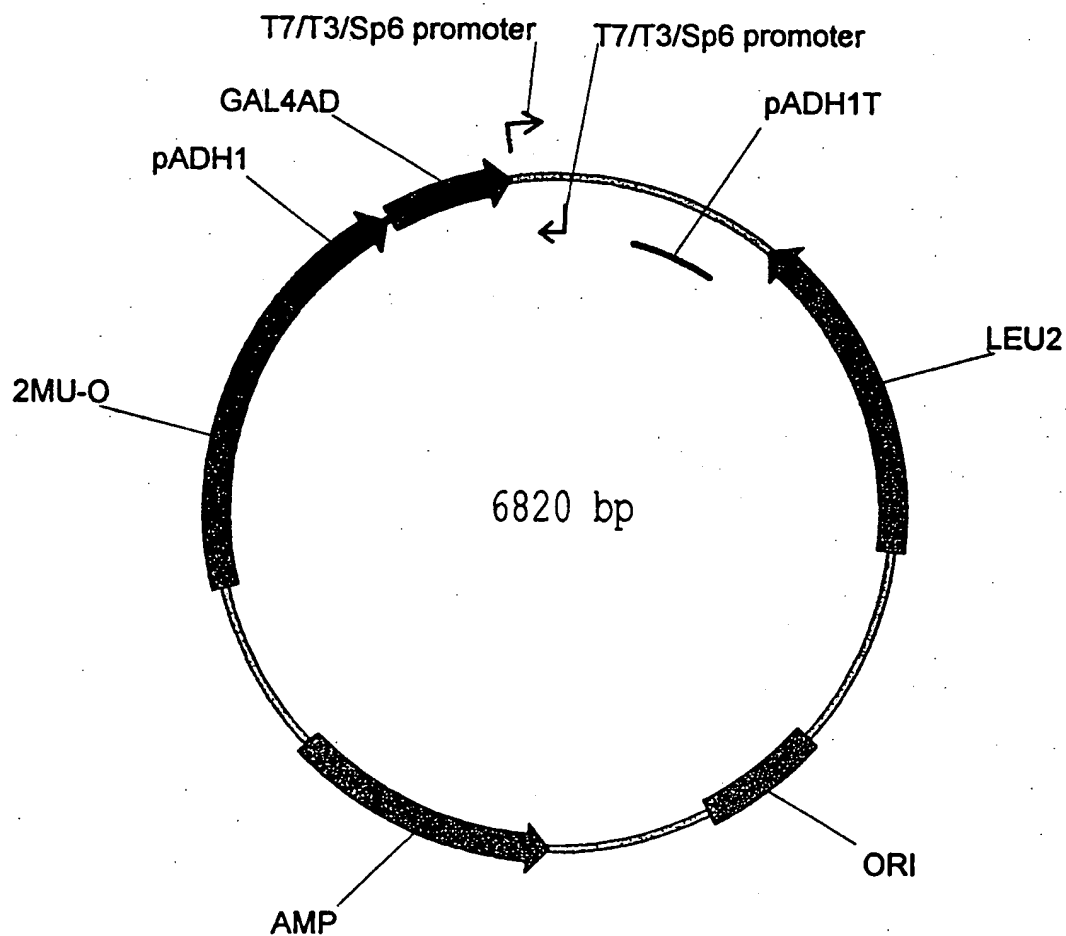
FIG. 14.



17/21

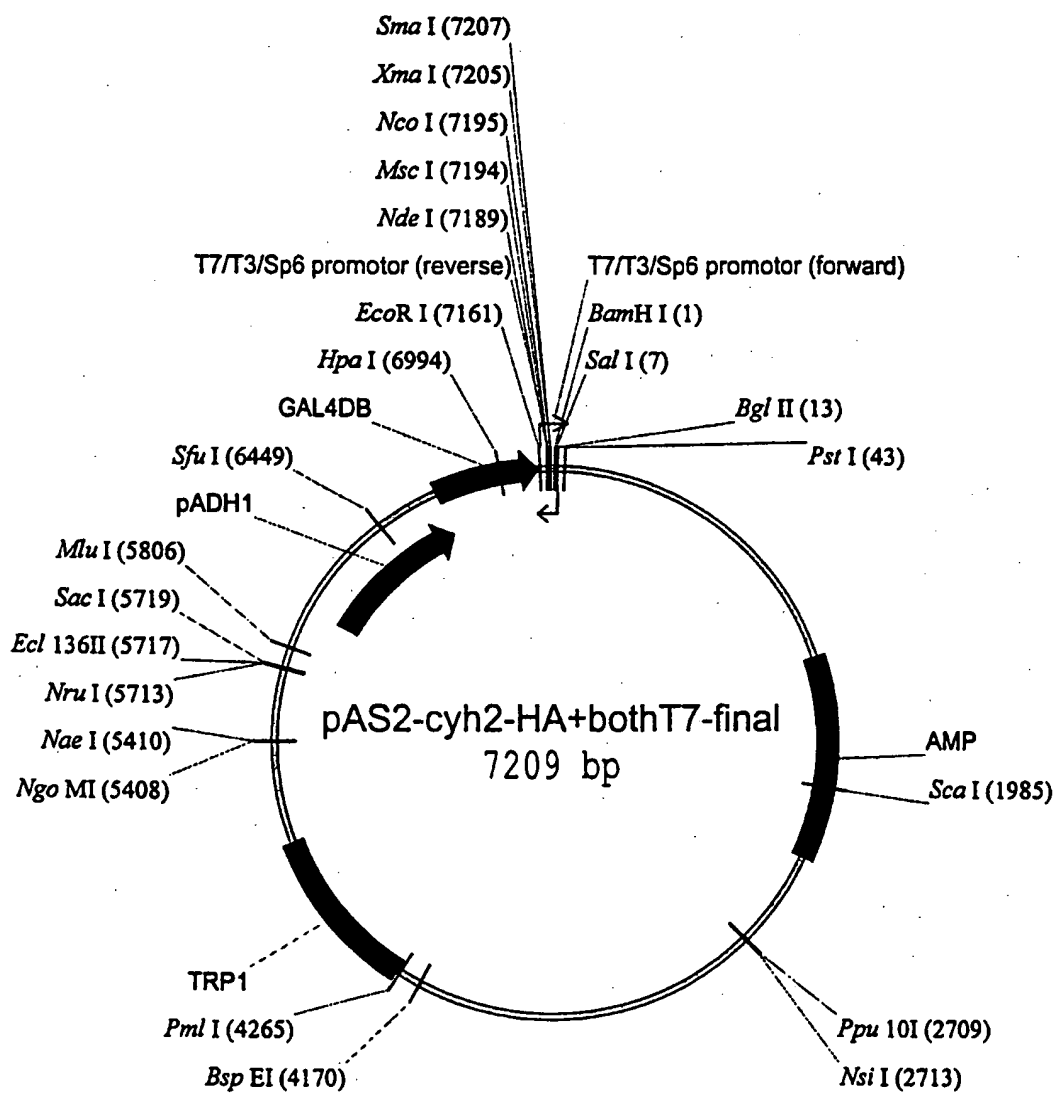
*FIG. 15.***pAS2* with Forward and Reverse T7/T3/Sp6**

18/21

*FIG. 16.***pGAD424 with Forward and Reverse T7/T3/Sp6**

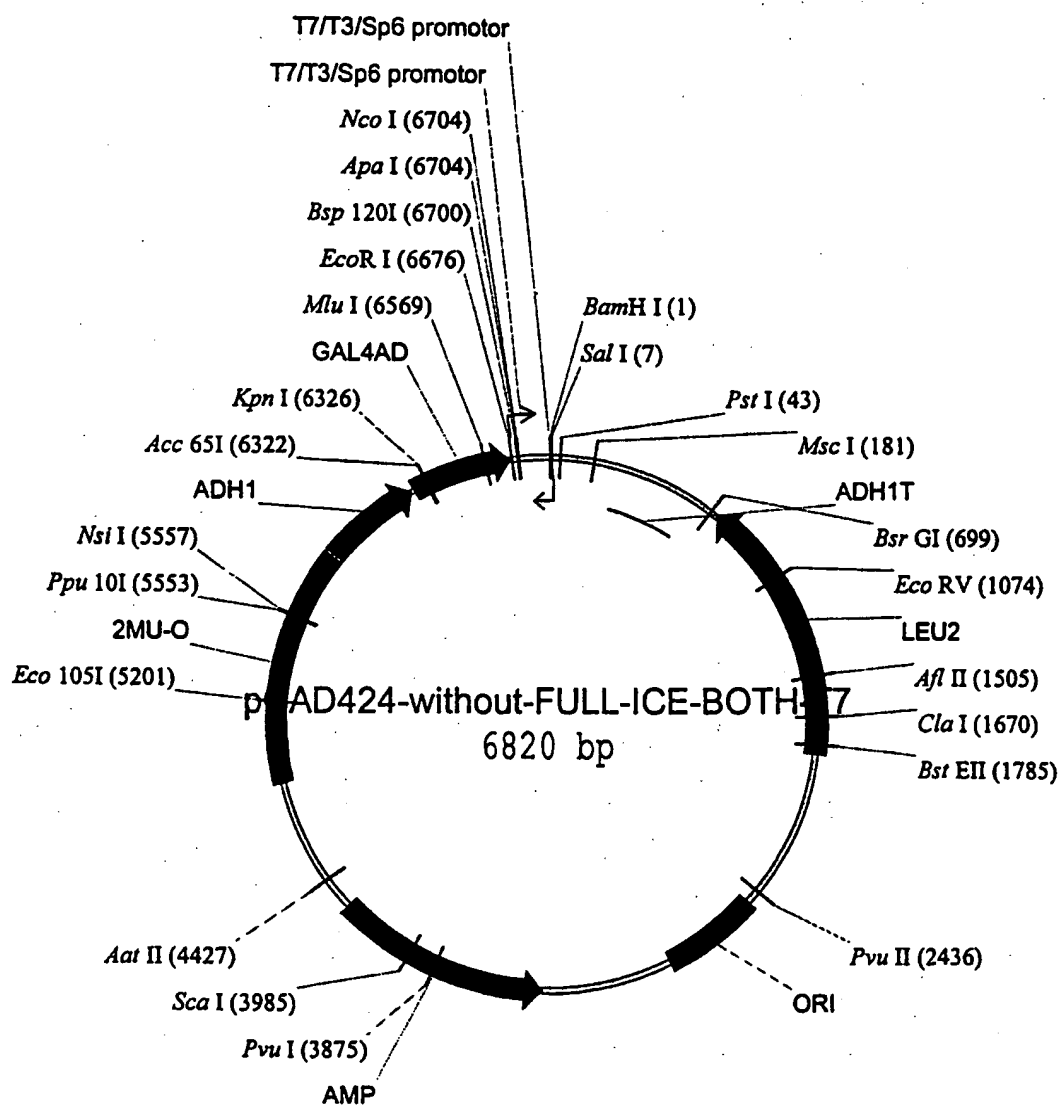
19/21

FIG. 17.



20/21

FIG. 18.



21/21

FIG. 19

